FACULTY OF ENGINEERING AND ARCHITECTURE

Diagnosis of depression based on resting state functional MRI

Gert Vanhollebeke
Student number: 01307496

Supervisors: Prof. dr. ir. Pieter van Mierlo, Prof. Chris Baeken
Counsellor: Prof. dr. ir. Pieter van Mierlo

Master's dissertation submitted in order to obtain the academic degree of
Master of Science in Biomedical Engineering

Academic year 2018-2019

GHENT UNIVERSITY

Acknowledgements

If a man has lost a leg or an eye, he knows he has lost a leg or an eye; but if he has lost a self—himself—he cannot know it, because he is no longer there to know it.

Oliver Sacks

As working on and writing a master's dissertation is no easy task, it is only normal to thank the people that helped me.

First of all I would like to thank my promotors, Prof. dr. ir. Pieter van Mierlo and Prof. Chris Baeken. The flexibility and willingness they both showed, that made it possible for me to explore and work out my own idea as well as their guidance and help throughout the year made the whole dissertation process much more pleasing.

Next, I would like to thank my parents. Living together with a student, staying calm when that student is annoying must be hard. My deepest thanks for the continuous support throughout the last year, and all years before this one. Special thanks to my father and sister for taking the time to read and correct this dissertation.

Thanks also to Debby Klooster and Toon Van de Maele for helping me with the structural features part of this dissertation.

Lastly I would like to thank my fellow students and friends who struggled together with me. Trouble shared is trouble halved, let's have a beer when this all is over.

Permission for usage

Gert Vanhollebeke, Ghent, May 2019

Diagnosis of depression based on resting state functional MRI

by
GERT VANHOLLEBEKE

Master's dissertation submitted in order to obtain the academic degree of
Master of Science in Biomedical Engineering

Academic year 2018 - 2019

Supervisors: Prof. dr. ir. PIETER VAN MIERLO, Prof. CHRIS BAEKEN

Counsellor: Prof. dr. ir. PIETER VAN MIERLO

Department of Electronics and Information systems
Chair: Prof. dr. ir. KOEN DE BOSSCHERE
Faculty of Engineering and Architecture
Ghent University

Abstract

While research uncovers new insights in the pathology of depression and defines new brain regions associated with the disease, the diagnosis of depression still remains a challenging task. Because of limited availability of psychologists and psychiatrists there is a large diagnosis delay. This leaves some patients in need of help, which can lead to the deterioration of the patient's mental health. In this master's dissertation, both structural MRI and resting state functional MRI data from healthy controls and patients with depression are used to obtain features in order to train a classifier capable of diagnosing depression. Three different feature types - MRI volumetry, fMRI intensity and fMRI functional connectivity - are used as input for the classifier. Good results, over 90% accuracy, were obtained using fMRI functional connectivity features and the combination of fMRI intensity and fMRI functional connectivity features resulted in accuracies of 94%, indicating that resting state functional MRI data can be used to reliably diagnose depression.

Keywords

Depression, Computer-aided diagnosis, MRI, Resting state fMRI, Machine learning

Diagnosis of depression based on resting state functional MRI

Gert Vanhollebeke
Supervisors: Prof. dr. ir. Pieter van Mierlo, Prof. Chris Baeken

***Abstract*—While research uncovers new insights in the pathology of depression and defines new brain regions associated with the disease, the diagnosis of depression still remains a challenging task. Because of limited availability of psychologists and psychiatrists there is a large diagnosis delay. This leaves some patients in need of help, which can lead to the deterioration of the patient's mental health. A computer-aided diagnosis tool using both structural MRI and resting state functional MRI data is developed. Features from healthy controls and patients with depression are used to obtain features in order to train a classifier capable of diagnosing depression. Three different feature types - MRI volumetry, fMRI intensity and fMRI functional connectivity - are used as input for the classifier. Good results, over 90% accuracy, were obtained using fMRI functional connectivity features and the combination of fMRI intensity and fMRI functional connectivity features resulted in accuracies of 94%, indicating that resting state functional MRI data can be used to reliably diagnose depression.**

***Index Terms*—Depression, Computer-aided diagnosis, MRI, Resting state fMRI, Machine learning**

I. Introduction

Depression is a common mental disorder resulting in a persistent saddened mood and anhedonia, possibly accompanied with other symptoms. Over 300 million people suffer from depression worldwide, making it one of the most common mental illnesses [1]. Depression is not a single disease, but a general name describing a multitude of symptoms. Many different causes have been defined.

The main method of diagnosing depression is a diagnostic interview in which a professional psychologist or psychiatrist examines the patient to understand the symptoms the patient experiences and asses the severity of the disease. This diagnosis method is based on symptoms, which do not always reflect the origin of the disease or any comorbidities that are present. Many countries also do not have enough mental health professionals, resulting in long waiting times for patients and inadequate care. Additional diagnosis tools capable of diagnosing depression reliably and fast while also capable of diagnosing depression subtypes such as medication- and treatment-resistant depression are needed.

Neuroimaging has proven to be useful in the diagnosis of neurological disorders such as epilepsy and multiple sclerosis, but is not yet used consistently in the diagnosis of mental disorders [2], [3]. Diagnosis tools for depression and other mental illnesses based on neuroimaging techniques that achieve high accuracies have been developed, but the features used for classification do not reflect aspects of brain anatomy and function that could be affected by depression, making the clinical validation of such diagnosis tools difficult.

In this paper, a computer-aided diagnosis tool based on anatomical MRI and resting state functional MRI (fMRI) scans from a data set of 106 people (60 healthy controls, 46 depression patients) capable of diagnosing depression has been developed. Section II describes the preprocessing steps, section III describes the different feature types and subtypes and defines the feature selection process.The features that have been selected can easily be linked to different aspects of brain anatomy and function, increasing the clinical value of the diagnosis tool. Section IV describes the clinical relevance of the found features, section V the classification training pipeline. Section VI describes the results and discussion, section VIII delineates the final conclusion that has been reached.

II. Feature preprocessing

All fMRI data is preprocessed using the CONN toolbox, which uses functions from the statistical parametric mapping software toolbox [4], [5]. A preprocessing pipeline is selected and adjusted to the needs of the data set.

All fMRI files are converted to the nifti file format, the first and last five scans are removed for signal equilibrium and signal dropout. Motion correction, slice timing correction and coregistration are applied. Afterwards the time series are high pass filtered to remove scanner drift; this is done using the MATLAB toolbox from Anthony Barone [6]. Finally the data is normalized to the MNI space and is smoothed using a gaussian kernel with 6mm width.

III. Feature selection

Three different feature types are investigated: intensity features, connectivity features and structural features. Each feature type reflects a different aspect of the brain and could possibly show alterations in brain anatomy and function due to depression. Each feature type contains multiple feature subtypes.

The feature selection process consists of three parts: a feature calculation process, assessing group differences and a feature selection process. In the feature calculation process all possible features of each type and subtype are calculated

for each person in the data set separately. In the assessment of group differences the feature values are averaged for each group (healthy controls and depression group) separately and the average depression group feature values are subtracted from the average healthy controls feature values. In the feature selection process, the twenty features with the highest difference between the average group feature values of both groups are selected as final features. The final features are tested using a two-pair t test to investigate their statistical significance. Finally the features are normalized using the z-score.

A. Structural features

Two different subtypes of structural features are defined: cortical thickness and brain parcel volume. These features, that are obtained using the FreeSurfer software [9], reflect possible anatomical changes that are present in people with depression. Contrary to the intensity and connectivity features, no feature selection process is used as fewer possible features are available. All possible features are tested using a two-pair t test for statistical significance. Only 19 left hemisphere cortical thickness, 11 right hemisphere cortical thickness and 6 parcel volume features are statistically significant.

B. Intensity features

Intensity features reflect the average activity of the brain through time. These features are obtained from fMRI data and could be compared (to a certain extent) to positron emission tomography (PET) or single photon emission computed tomography (SPECT) imaging as they show metabolic processes such as glucose uptake within the body and brain. Two different feature subtypes are defined: absolute and relative intensity features.

1) Absolute intensity features: Absolute intensity features show the absolute activity in the brain and are calculated by averaging all time series belonging to a certain brain parcel, resulting in an average activity of each brain parcel. These features are prone to differences between patient scans as they are not normalized. Global elevations or decreases in intensity between patient scans have a large influence on the feature values.

2) Relative intensity features: Relative intensity features show the relative activity in the brain and are calculated by averaging all time series belonging to a brain parcel to a single value (similar to absolute intensity features). The average activity value of each brain parcel is normalized using formula 1 where $I_{rel,j}$ is the relative intensity value of brain parcel j, $I_{abs,j}$ is the absolute intensity value of brain parcel j and $I_{brain,avg}$ is the average intensity value of the whole brain, calculated by averaging all averaged time series in the brain. This subtype is more resilient against global variations of intensity between patient scans.

$$I_{rel,j} = \frac{I_{abs,j} - I_{brain,avg}}{I_{brain,avg}} \tag{1}$$

C. Functional connectivity features

Connectivity features reflect the possible changes in functional connectivity of the brain due to depression. These features are obtained from the MRI data. Two connectivity measures are used: correlation and mutual information [7]. These connectivity measures are chosen as they are undirected and reflect functional connectivity in the time domain. The influence of global signal regression on classification is also investigated. As global signal regression is subject to much discussion the connectivity measures are calculated on both a global signal regressed data set and a non-regressed data set [8]. No extensions, such as graph based features, are explored from this feature type as the interpretability and clinical relevance of these types of features are low.

IV. Clinical relevance assessment

The features are obtained using a data-driven approach, no knowledge about depression is used as a bias in the feature selection process. The clinical relevance of the selected features is therefore not certain and is assessed. The structural features are first discussed. The intensity features will be discussed as one group as most features from both subtypes are the same. Only the connectivity features calculated on the regressed data set will be discussed as they resulted in much higher classification accuracies.

A. Structural features

1) Cortical thickness: A total of 30 cortical thickness features are statistically significant. Nineteen features are located in the left hemisphere, eleven in the right hemisphere. In the people in the depression group, all cortical regions have a decreased thickness, possibly pointing to neuronal atrophy due to depression. Several cortical regions that are statistically significant are related to depression: the left hemisphere rostral middle frontal gyrus, precentral gyrus, insula, precuneus, pars orbitalis, frontal pole, superior frontal gyrus, post central gyrus, caudal middle frontal gyrus, the right hemisphere frontal pole and superior frontal gyrus [15], [16].

2) Parcel volume: Six parcel volume features are statistically significant. Both the left and right cerebellum hemisphere and left and right caudate nucleus are found, which are linked to depression [17], [18].

B. Intensity features

Most of the features (absolute intensity features: 16 out of 20 features, relative intensity features: 16 out of 20 features) lie within four brain regions: the left and right superior frontal gyrus and the left and right rostral middle frontal gyrus. All regions lie within the prefrontal cortex and are linked to depression. All but two features show less activity in the depression group when compared to the healthy controls, reflecting the hypoactivity commonly found in people with depression [10].

C. Connectivity features

1) Correlation: Some features are possibly linked to brain parcels afflicted by depression such as the connection between the right anterior cingulate cortex and the left orbitofrontal cortex and the connection between the left and right prefrontal cortex [11]. Connections involving the precuneus could also be related to depression as the precuneus is part of the default mode network: disturbances in this network have been found in patients with depression [12].

2) Mutual information: Four brain regions are present in almost all mutual information features: the right paracentral gyrus, the right inferior parietal gyrus, the right superior parietal gyrus and the brain stem. Both increased and decreased functional connectivity have been found in the paracentral gyrus of people with depression, but its specific role in the disease is not yet known [13]. Contrary to literature, which reports reduced functional connectivity in both the right inferior and superior parietal gyrus, increased functional connectivity is found in this data set [14]. The brain stem contains several nuclei that possibly are involved in depression, but the specific connections found here (connections between the brain stem and the right inferior parietal gyrus, the right superior parietal gyrus and the right paracentral gyrus) have not been described in literature.

V. Classifier training pipeline

To be able to compare the results of different classifiers, a classifier training pipeline is defined and is shown in figure 1. A first selection of 46 people from the healthy controls is made to avoid class imbalance when training a classifier. Secondly a train and test set is defined from the selected individuals. The 80-20 rule is used, resulting in a validation set of 19 people and a training set of 73 people. The model is trained and afterwards validated. Both the first selection and the train-test selection is performed using a random permutation. This ensures that each time the pipeline is used, a different variation of the available data is used. Important to notice is that two validation sets are used to validate a trained classifier. The official validation set, containing nineteen people from both the healthy controls and the depression group, and an optional validation set, containing fourteen people, all healthy controls.

VI. Results

Two different categories of classifiers have been trained: single feature and combined feature classifiers. Single feature classifiers are trained on a feature set containing a single feature subtype. Combined feature classifiers are trained on a feature set containing two or three subtypes belonging to one feature type. Most results of the combined feature classifiers are lower than the single feature classifiers, the exception to this is a classifier trained on de combined feature sets of the absolute intensity, relative intensity, correlation with regressed data and mutual information with regressed data features. Each single feature classifier is trained sixty times in total. Three different amounts of features (variable amount, 10 features and 20 features) are used as input for the classifier and each feature amount is trained twenty times. This leads to a result distribution of twenty samples for each feature amount. The final results are shown as the average accuracy, calculated from the best performing result distribution of the official validation set and the corresponding result distribution of the optional validation set. The average standard deviation (SD) of these validation sets is also shown. It is a measure for the variability of the result distributions: the higher the standard deviation, the lower the reliability of the feature set.

A. Single feature classifiers

Nine different single feature classifiers are trained using the nine different feature subtypes. The results are shown in table I. Several conclusions can be formed from the results. Firstly the results show that both the intensity features have a comparable accuracy and SD, showing that the absolute intensity features did not suffer from any possible global variation of intensity between patients. Secondly the results show that global signal regression has a significant positive influence on the connectivity features. An increase in accuracy of $\pm 6\%$ for the correlation features and $\pm 30\%$ for the mutual information features is obtained when global signal regression is used. Decreases in SD of both feature types also show the increased quality of the features. Thirdly the results show that the structural features are not capable of accurately classifying depression. A reason for this could be the assumption that all people in the depression group have suffered the same form of depression, have taken the same medications and underwent the same therapies, while this is in reality not true. Outliers were found in the intensity and connectivity classifiers. A classifier with an accuracy of 90.9% is obtained (not shown in table I as this shows the mean accuracy) in both the correlation and mutual information feature sets (calculated with the regressed data set).

TABLE I
Mean results of the single feature classifiers.

Feature type	Acc. (%)	Sens.	Spec.	PPV	NPV
LH thickness	59,6 ± 15,1	0.58	0.525	0.55	0.556
RH thickness	56,3 ± 14	0.626	0.574	0.595	0.605
Parcel volume	61 ± 20	0.558	0.504	0.535	0.528
Abs. Int.	74,7 ± 10,5	0.762	0.718	0.72	0.75
Rel. Int.	73,3 ± 11,3	0.761	0.714	0.73	0.744
Corr. non regr.	77,5 ± 15,8	0.803	0.756	0.77	0.789
Mut. Inf. non regr.	49,5 ± 19,9	0.485	0.431	0.46	0.456
Corr. regr.	**83,1 ± 7,3**	**0.858**	**0.807**	**0.815**	**0.85**
Mut. Inf. regr.	79,2 ± 7,5	0.768	0.817	0.805	0.78

B. Combined feature classifiers

Four different combined feature classifiers are trained: an intensity feature classifier, a connectivity feature classifier, a structural feature classifier and a intensity and connectivity classifier. The intensity classifier is trained using a feature set that contains both the absolute and relative intensity features. The connectivity feature classifier is trained using a feature set that contains both the correlation and mutual information features calculated from the regressed data set. The structural feature classifier is trained using all structural features. The intensity and connectivity classifier is trained using the

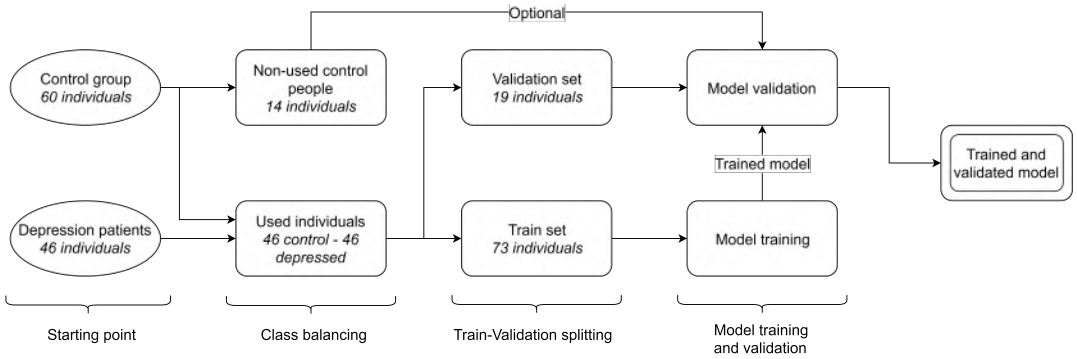


Fig. 1. Visualization of the classification pipeline.

absolute intensity features, the relative intensity features, the correlation with regressed data features and the mutual information with regressed data features. These classifiers are only trained twenty times with all features. The results are shown in table II. Multiple conclusions can be made from the results. Firstly the results show that most combined feature classifiers perform worse than the single feature classifiers. This is counter-intuitive as more features would normally result in a better classification. A possible explanation for this could be that each feature contains some unwanted noise and that more features result in more unwanted noise, reducing the classification accuracy. Secondly the results show that the reduction in classification accuracy, when compared to the results of their respective subtypes, is much higher for the connectivity feature classifier than for the other combined classifiers (the structural classifier even has a small increase in accuracy). This is again counter-intuitive as the single type classifiers using the same features have the highest accuracy.

The intensity and connectivity classifier has, contrary to the other combined feature classifiers, a higher accuracy than the single feature classifiers. A classifier from this type is obtained with an accuracy of $\pm$94.7% (not shown in table II as this shows the mean accuracy). The increase in accuracy compared to both the single feature and other combined feature classifiers is explained by the fact that more features are used (80 features). The classifier is trained with the data of 73 people (see section V), which is less than the amount of features used. SVMs are able to correctly train with more features than samples, but are prone to overfitting. Further validation is needed for this classifier type.

TABLE II
RESULTS OF THE COMBINED FEATURE CLASSIFIERS.

Feature type	Acc. (%)	Sens.	Spec.	PPV	NPV
Structural	61,5 $\pm$ 12,4	0.613	0.561	0.58	0.594
Intensity	70,3 $\pm$ 14	0.738	0.684	0.695	0.728
Connectivity	61,8 $\pm$ 13,7	0.688	0.629	0.645	0.672
Int. and Conn.	**88.7 $\pm$ 6.97**	**0.879**	**0.827**	**0.835**	**0.872**

VII. DISCUSSION

The obtained results (best mean accuracy = $\pm$88%, best accuracy = $\pm$94.7%) are comparable or higher to those found in literature [19], [20] [21]. The clinical relevance as well as the easy interpretation of the used features, which is not found in literature, makes the classifiers that are obtained highly relevant. Structural features can not yet be used for classification, but the addition of information about the severity and duration of the depressive episodes of the patients could solve this problem. Both intensity and functional connectivity features prove to be adequate for classification. While combining feature subtypes from a single feature type did not lead to an increase in accuracy, the combination of both intensity and functional connectivity features resulted in the best performing classifiers.

VIII. CONCLUSION

A computer-aided diagnosis tool has been developed using both anatomical and fMRI data to diagnose depression. A data-driven approach without any prior assumptions or knowledge about the known effects of depression has been used to select features for classification. The clinical relevance of the features has been assessed and all feature types contain features that can be linked to depression from a clinical point of view. This shows that, while using a different approach, the same brain regions that are influenced and change due to depression are identified. High classification accuracies have been obtained when the features are used to distinguish depression from healthy controls. The highest accuracies are obtained when the intensity and connectivity features are combined. As a general conclusion, it can be stated that resting state fMRI data can be used to accurately predict depression and could be used in the future to aid mental health professionals for fast and reliable diagnoses thus reducing their workloads while also reducing waiting times for patients.

REFERENCES

[1] World Health Organization. (2017). Depression and other common mental disorders: global health estimates (No. WHO/MSD/MER/2017.2). World Health Organization.

[2] Olson, L. D., & Perry, M. S. (2013). Localization of epileptic foci using multimodality neuroimaging. International journal of neural systems, 23(01), 1230001.

[3] Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., ... & Lublin, F. D. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of neurology, 69(2), 292-302.

[4] Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain connectivity, 2(3), 125-141.
[5] Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (Eds.). (2011). Statistical parametric mapping: the analysis of functional brain images. Elsevier.
[6] "myfreqfilter", Anthony Barone, The University of Texas at Austin, Institute for Geophysics.
[7] Cover, T. M., & Thomas, J. A. (2012). Elements of information theory. John Wiley & Sons.
[8] Saad, Z. S., Gotts, S. J., Murphy, K., Chen, G., Jo, H. J., Martin, A., and Cox, R. W. (2012). Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. Brain connectivity, 2(1), 25-32.
[9] Fischl, B. (2012). FreeSurfer. Neuroimage, 62(2), 774-781.
[10] George, M. S., Ketter, T. A., & Post, R. M. (1994). Prefrontal cortex dysfunction in clinical depression. Depression, 2(2), 59-72.
[11] Ramirez-Mahaluf, J. P., Perramon, J., Otal, B., Villoslada, P., & Compte, A. (2018). Subgenual anterior cingulate cortex controls sadness-induced modulations of cognitive and emotional network hubs. Scientific reports, 8(1), 8566.
[12] Utevsky, A. V., Smith, D. V., & Huettel, S. A. (2014). Precuneus is a functional core of the default-mode network. Journal of Neuroscience, 34(3), 932-940.
[13] Kenny, E. R., O'brien, J. T., Cousins, D. A., Richardson, J., Thomas, A. J., Firbank, M. J., & Blamire, A. M. (2010). Functional connectivity in late-life depression using resting-state functional magnetic resonance imaging. The American Journal of Geriatric Psychiatry, 18(7), 643-651.
[14] Dutta, A., McKie, S., & Deakin, J. W. (2014). Resting state networks in major depressive disorder. Psychiatry Research: Neuroimaging, 224(3), 139-151.
[15] Niu, M., Wang, Y., Jia, Y., Wang, J., Zhong, S., Lin, J., ... & Huang, R. (2017). Common and specific abnormalities in cortical thickness in patients with major depressive and bipolar disorders. EBioMedicine, 16, 162-171.
[16] Mackin, R. S., Tosun, D., Mueller, S. G., Lee, J. Y., Insel, P., Schuff, N., ... & Weiner, M. W. (2013). Patterns of reduced cortical thickness in late-life depression and relationship to psychotherapeutic response. The American Journal of Geriatric Psychiatry, 21(8), 794-802.
[17] Baldaçara, L., Borgio, J. G. F., Lacerda, A. L. T. D., & Jackowski, A. P. (2008). Cerebellum and psychiatric disorders. Brazilian Journal of Psychiatry, 30(3), 281-289.
[18] Kim, M. J., Hamilton, J. P., & Gotlib, I. H. (2008). Reduced caudate gray matter volume in women with major depressive disorder. Psychiatry Research: Neuroimaging, 164(2), 114-122.
[19] J. R. Sato, J. Moll, S. Green, J. F. Deakin, C. E. Thomaz, and R. Zahn, "Machine learning algorithm accurately detects fmri signature of vulnerability to major depression," Psychiatry Research: Neuroimaging 233, 289 (2015).
[20] M. J. Patel, A. Khalaf, and H. J. Aizenstein, "Studying depression using imaging and machine learning methods," NeuroImage: Clinical 10, 115 (2016).
[21] M. Wei, J. Qin, R. Yan, H. Li, Z. Yao, and Q. Lu, "Identifying major depressive disorder using hurst exponent of resting-state brain networks," Psychiatry Research: Neuroimaging 214, 306 (2013).

Contents

Acknowledgements i

Permission for usage iii

Abstract iv

Extended abstract vii

Table Of Contents xii

List of Figures xix

List of Tables xxi

1 Introduction 1

1.1 Problem definition . . . 1

1.2 Objective . . . 4

2 MRI and fMRI 7

2.1 MRI . . . 7

2.1.1 Physical principles . . . 7

2.1.1.1 Magnetic moment . . . 7

2.1.1.2 Total nuclear spin . . . 8

2.1.1.3 The larmor frequency . . . 8

2.1.1.4 Relevance for MRI imaging . . . 9

2.1.2 The MRI scanner . . . 10

2.1.2.1 The homogeneous magnetic field . . . 10

2.1.2.2 The rotating magnetic field . . . 11

2.1.2.3 The gradient magnetic field . . . 12

2.1.3 MRI imaging sequences . . . 12

2.1.3.1 Echo planar imaging . . . 12

2.2 Functional MRI . . . 13

2.2.1 The BOLD response . . . 13

2.2.2 Advantages and limitations of fMRI . . . 14
2.2.3 Resting state fMRI and task-related fMRI . . . 14

3 Machine learning **15**
3.1 Introduction . . . 15
3.1.1 Supervised learning . . . 15
3.2 Support vector machines . . . 15

4 Structural features **19**
4.1 Available dataset . . . 19
4.2 Structural features . . . 20
4.2.1 FreeSurfer . . . 20
4.2.1.1 Step 1: Normalization . . . 20
4.2.1.2 Step 2: Skull strip . . . 20
4.2.1.3 Step 3: Subcortical segmentation . . . 20
4.2.1.4 Step 4: Statistics calculation . . . 20
4.2.1.5 Step 5: White matter segmentation . . . 21
4.2.1.6 Step 6: Brain division . . . 21
4.2.1.7 Step 7: Tesselation . . . 21
4.2.1.8 Step 8: Surface smoothing . . . 21
4.2.1.9 Step 9: Inflation . . . 21
4.2.1.10 Step 10: Surface definitions . . . 21
4.2.1.11 Step 11: Spherical inflation . . . 21
4.2.1.12 Step 12: Parcel labeling and statistics calculation . . . 21
4.2.2 Interpretation of the found features . . . 22
4.2.2.1 Left hemisphere thickness . . . 22
4.2.2.2 Right hemisphere thickness . . . 23
4.2.2.3 Parcel volume . . . 23

5 Functional features **25**
5.1 Preprocessing process . . . 25
5.1.1 Step 0: File conversion . . . 26
5.1.2 Step 1: Removal of the first and last scans for signal equilibrium . . . 26
5.1.3 Step 2: Motion correction . . . 26
5.1.4 Step 3: Slice timing correction . . . 27
5.1.5 Step 4: Coregistration . . . 27
5.1.6 Step 5: High pass filtering . . . 27
5.1.7 Step 6: Normalization . . . 28
5.1.8 Step 7: Spatial smoothing . . . 29
5.2 Parcellation . . . 30

5.2.1 Human brain atlas . . . 30
5.2.2 Human brain atlas resizing . . . 30
5.3 Intensity features . . . 30
5.3.1 Absolute intensity and relative intensity . . . 30
5.3.2 Feature selection process . . . 31
5.3.2.1 Time averaging . . . 31
5.3.2.2 Parcel averaging . . . 31
5.3.2.3 Group averaging . . . 32
5.3.2.4 Difference calculation . . . 32
5.3.2.5 Feature selection . . . 32
5.3.3 Interpretation of the found features . . . 33
5.3.3.1 Interpretation of the used tables . . . 33
5.3.3.2 Absolute intensity features . . . 33
5.3.3.3 Relative intensity features . . . 34
5.4 Connectivity features . . . 36
5.4.1 Functional connectivity measures . . . 36
5.4.1.1 Selection process of connectivity measures . . . 37
5.4.2 Feature selection process . . . 38
5.4.2.1 Global signal regression . . . 38
5.4.2.2 Parcel simplification . . . 39
5.4.2.3 Connectivity measure calculation . . . 40
5.4.2.4 Group averaging . . . 40
5.4.2.5 Difference calculation . . . 40
5.4.2.6 Feature selection . . . 40
5.4.3 Interpretation of the found features . . . 41
5.4.3.1 Features calculated with the non-regressed data set . . . 41
5.4.3.2 Features calculated with the regressed data set . . . 42

6 Classifier training **45**
6.1 Classification training pipeline . . . 45
6.1.1 Starting point . . . 45
6.1.2 Class balancing . . . 45
6.1.3 Train-validation splitting . . . 46
6.1.4 Model training and validation . . . 46
6.1.4.1 Training . . . 46
6.1.4.2 Validation . . . 46

7 Results **47**
7.1 Data collection and presentation . . . 47
7.1.1 Data collection . . . 47
7.1.2 Presentation . . . 47
7.1.2.1 Violin plot . . . 48
7.2 Global results . . . 49
7.3 Results . . . 50
7.3.1 Single feature type classifier . . . 50
7.3.1.1 The left hemisphere thickness classifier . . . 50
7.3.1.2 The right hemisphere thickness classifier . . . 51
7.3.1.3 The parcel volume classifier . . . 52
7.3.1.4 The absolute intensity classifier . . . 54
7.3.1.5 The relative intensity classifier . . . 55
7.3.1.6 The correlation with non-regressed data classifier . . . 56
7.3.1.7 The mutual information with non-regressed data classifier . . . 58
7.3.1.8 The correlation with regressed data classifier . . . 59
7.3.1.9 The mutual information with regressed data classifier . . . 61
7.3.2 Combined-feature classifiers . . . 62
7.3.2.1 The structural feature classifier . . . 62
7.3.2.2 The intensity feature classifier . . . 63
7.3.2.3 The connectivity feature classifier . . . 65
7.3.2.4 The intensity and connectivity feature classifier . . . 66

8 Discussion **69**
8.1 Part 1: Feature type specific . . . 69
8.1.1 Intensity features . . . 69
8.1.2 Connectivity features . . . 69
8.1.3 Structural features . . . 70
8.2 Part 2: Feature type comparison . . . 70
8.3 Part 3: Combined feature classifiers . . . 71
8.3.1 The intensity feature classifier . . . 71
8.3.2 The connectivity feature classifier . . . 71
8.3.3 The structural feature classifier . . . 71
8.3.4 The intensity and connectivity feature classifier . . . 71
8.4 Part 4: Performance with respect to atlas level . . . 72

9 Conclusion **74**

Bibliography **76**

Appendices **82**

A MRI parameters of the fMRI data **83**

List of Figures

1.1 Prevalence of depressive disorders in the world [1]. 2

1.2 Prevalence of depressive disorders with respect to age [1]. 2

2.1 Visualization of the angle (ϕ) between the magnetic moment vector (μ) of a nucleus (shown in red) and an external magnetic field (B) and the corresponding precession.[10] 8

2.2 Example of an MRI image. The different tissues (grey matter, white matter, skin, air) can clearly be distinguished. 9

2.3 Simplified representation of an MRI scanner [2]. 10

2.4 Random orientation of the hydrogen atoms without an external magnetic field (left), net magnetization M_L due to a magnetic field B_0 (right). 10

2.5 Visualization of the two components of the net magnetization M. 11

2.6 Visualization of the longitudinal (T1) and transverse (T2) recovery [3]. 11

2.7 Visualization of the BOLD response with respect to time. The x-axis shows time, the y-axis the change in oxy-/deoxyhemoglobin concentration in the blood in percentage. 14

3.1 A possible distribution of data points. Data points beloning to C_1 are shown by red dots, C_2 is shown by green dots. V_1 and V_2 denote two variables. 16

3.2 Visual presentation of the optimal hyperspace separating C_1 and C_2. The support vectors are encircled. 17

5.1 The complete preprocessing pipeline. 25

5.2 Visual representation of the T1 relaxation of the tissue and the corresponding BOLD signal. The vertical lines show the start of a new scan. (based on [4]) 26

5.3 Visualization of the initial presence of scanner drift (a) and the result after high pass filtering (b). 28

5.4 Visualization of the effect of normalization. 29

5.5 Visualization of the effect of spatial smoothing. 29

5.6 The intensity feature selection process. 31

5.7 The connectivity feature selection process. 38

6.1 The complete classification training pipeline. 46

7.1 Example of the different possible results and violin plots. 48

7.2 Violin plot of the global results. 49

7.3 Violin plot of the results of the left hemisphere thickness classifier. 50
7.4 Violin plot of the results of the right hemisphere thickness classifier. 52
7.5 Violin plot of the results of the parcel volume classifier. 53
7.6 Violin plot of the results of the absolute intensity classifier. 55
7.7 Violin plot of the results of the relative intensity classifier. 56
7.8 Violin plot of the results of the correlation with non-regressed data classifier. 57
7.9 Violin plot of the results of the mutual information with non-regressed data classifier. 58
7.10 Violin plot of the results of the correlation with regressed data classifier. 60
7.11 Violin plot of the results of the mutual information with regressed data classifier. . . . 62
7.12 Violin plot of the results of the structural feature classifier. 63
7.13 Violin plot of the results of the intensity feature classifier. 64
7.14 Violin plot of the results of the connectivity feature classifier. 65
7.15 Violin plot of the results of the connectivity feature classifier. 66

8.1 Violin plot of the results of the connectivity feature classifier. 72

List of Tables

4.1 Significant left hemisphere thickness features. 22
4.2 Significant right hemisphere thickness features. 23
4.3 Significant parcel volume features. 23

5.1 Absolute intensity features (Atlas3). 34
5.2 Relative intensity features (Atlas5). 35
5.3 Functional connectivity measures [4]. 36
5.4 Correlation with non-regressed data set features (Atlas2). 41
5.5 Mutual information with non-regressed data set features (Atlas3). 42
5.6 Correlation with regressed data set features (Atlas3). 43
5.7 Mutual information with regressed data set features (Atlas3). 44

7.1 Best results of the left hemisphere thickness classifier. 50
7.2 Best results of the right hemisphere thickness classifier. 51
7.3 Best results of the parcel volume classifier 53
7.4 Best results of the absolute intensity feature classifier (Atlas3). 54
7.5 Best results of the relative intensity feature classifier (Atlas5). 55
7.6 Best results of the correlation with non-regressed data classifier (Atlas2). 57
7.7 Best results of the mutual information with non-regressed data classifier (Atlas3). . . . 59
7.8 Best results of the correlation with regressed data classifier (Atlas3). 60
7.9 Best results of the mutual information with regressed data classifier (Atlas3). 61
7.10 Best results of the structural feature classifier 62
7.11 Best results of the intensity feature classifier 64
7.12 Best results of the connectivity feature classifier 65
7.13 Best results of the intensity and connectivity feature classifier. 67

Chapter 1

Introduction

1.1 Problem definition

The *World Health Organization* (WHO) defines depression as *"a common mental disorder, characterized by persistent sadness and a loss of interest in activities that you normally enjoy, accompanied by an inability to carry out daily activities, for at least two weeks"*.[1] It is estimated that more than 300 million people worldwide suffer from some form of depression, making it the most prevalent mental disorder.[2]

Depression is not a single disease having a single origin, but rather a group of illnesses that result in similar symptoms, varying both in duration and severity. The origin of depression is widespread, ranging from neurotransmitter imbalances [5], medication side effects [6], substance use and abuse [7], childhood trauma [8] to mood changes due to chronic illness [9]. Recent studies report the major influence of gut bacteria [10] and infection [11] as initiators of depression. Rapid changes in social, cultural and environmental aspects of the society are also blamed for the recent elevation in prevalence and incidence of depression [12], [13].

Depression is mainly treated with psychotherapy and medication. More severe forms, mainly medication-resistant depression, can be treated using neurostimulation methods like *transcranial magnetic stimulation* (TMS), *direct current stimulation* (DCS) [14] or *deep brain stimulation* (DBS). DBS and TMS are relatively new types of treatment and their efficacy as well as optimizations are still researched. Patients with very severe forms of depression that cannot (sufficiently) be treated with the previously mentioned methods, can sometimes be helped with *electroconvulsion therapy* (ECT) [15].

The two major symptoms of depression (as defined by the *Diagnostic and Statistical Manual of Mental Disorders, fifth edition* (DSM-5)) are a depressive mood and a lack of interest or pleasure in most activities that are present for at least two weeks [16]. It should be noted that it is not necessary for both symptoms to be present. Further symptoms such as weight change, irregular sleep patterns, motor abnormalities, increased feelings of guilt or worthlessness, decreased concentration and suicidal thoughts, ideations or attempts are defined and could help the differential diagnosis and assess the severity of the illness.

According to the WHO, depression is the leading cause of disability worldwide and a major contributor to the global burden of disease [1]. An estimated 5.1% of women and 3.6% of men, equaling around 322 million people worldwide, suffer from some form of depression. The prevalence of depressive disorders throughout the world varies and is shown in figure 1.1, the variation of depressive disorders prevalence

1. http://www.who.int/mental_health/management/depression/en/
2. https://www.who.int/news-room/fact-sheets/detail/depression

with respect to age is shown in figure 1.2.

In Belgium specifically, 700 000 people struggle with mental health issues yearly, a considerable amount of these people with depression [17], [18]. On average three people commit suicide in Belgium each day, making it one of the countries with the highest suicide rates in Europe. Belgium has the 11th highest suicide rate in the world.[3] Only one in three people experiencing mental health issues reach out for professional mental help, showing the broader problems around mental health such as stigma and lack of access to proper professional help.[4]

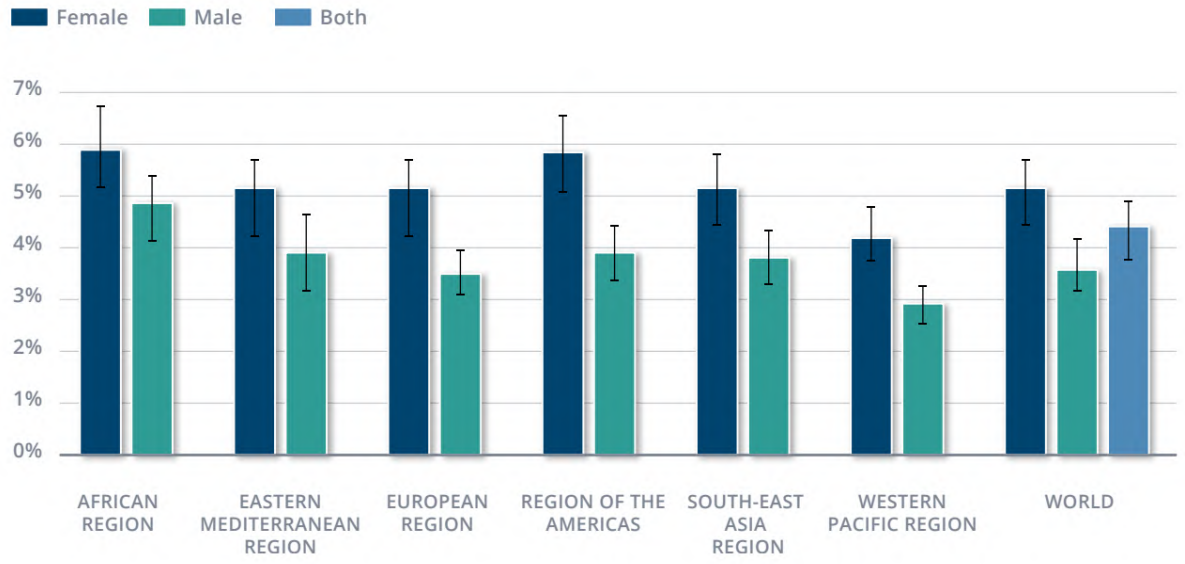


Figure 1.1: Prevalence of depressive disorders in the world [1].

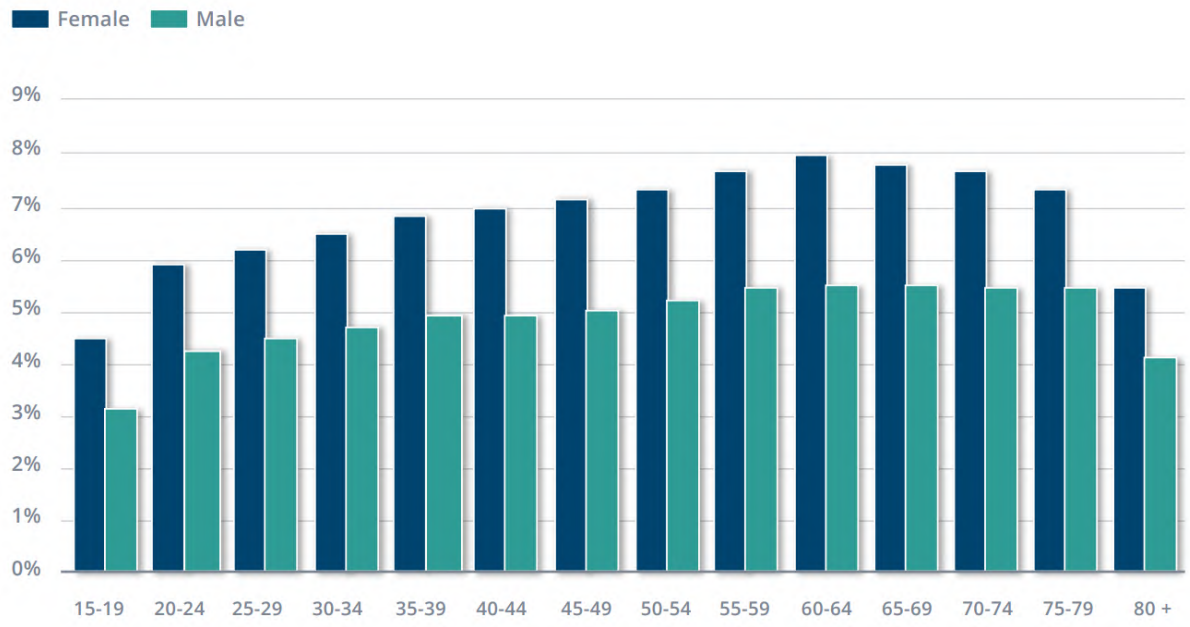


Figure 1.2: Prevalence of depressive disorders with respect to age [1].

3. http://worldpopulationreview.com/countries/suicide-rate-by-country/
4. https://www.geestelijkgezondvlaanderen.be/feiten-cijfers

Diagnosis of depression

The diagnosis of depression is done in a diagnostic interview such as the *structural clinical interview for DSM* (SCID) [19]. The duration of an interview varies based on the complexity of the disorder and the ability of the patient to correctly describe the symptoms and takes on average between one and two hours.[5]

Diagnostic interviews are the main diagnosis tool for mental disorders such as depression and have high accuracy when performed by experienced psychologists and psychiatrists [20]. Self-administered depression measures such as the *Patient Health Questionnaire-9* (PHQ-9) and the *Geriatric Depression Scale-15* (GDS-15) exist and could provide an useful first step for people experiencing mental health issues, but the accuracy of such tests is lower as people do not have the training and experience required to correctly recognize symptoms and severity and to distinguish between the different disorders that exist [21].

A diagnostic interview is limited by the fact that it is based on symptoms. The multiple origins as well as the possible resistances to certain treatment options (medication-resistant and treatment-resistant depression) do not always translate to variations in symptoms and are thus not recognizable in a diagnostic interview. Because of this, first treatment of depression is focused on trail periods of different medications until one is found that alleviates the symptoms adequately. This trail period can be considerable and can even be futile when a patient has medication-resistant depression.

While the diagnosis of depression can be accurate [22], the waiting time for a diagnostic interview can be long. The VVP (*Vlaamse vereniging voor psychiatrie*) recently mentioned the fact that waiting times for an appointment with a mental health professional can be as much as 18 months [17].[6] A person experiencing serious mental health problems is however often in urgent need of psychological or psychiatric guidance and a long waiting period may worsen the mental health of the person as well as the possible outcomes. Limited access to psychological and psychiatric care has been linked multiple times to the rate of suicide.[7]

The limitations of the diagnostic interview as well as the long waiting times for professional help give rise to a need for additional diagnostic tools. Neuroimaging techniques are promisins as a new diagnostic tool as they are routinely used to diagnose neurological diseases such as multiple sclerosis (MS) [8] and epilepsy [9] and could possibly be used to also diagnose depression. A major difference between the previously mentioned diseases and depression is that depression does not (yet) have a well defined origin within the brain (contrary to MS where brain lesions can be seen on a *magnetic resonance imaging* (MRI) scan) and has no easily measurable neurological symptoms (contrary to epilepsy where seizures can be recognized in electroencephalography (EEG) as drastic changes in amplitude and frequency). This makes the use of neuroimaging to diagnose depression more reliant on complex algorithms than on visual analysis of the physician.

Several imaging techniques, such as EEG and functional MRI (fMRI), have already been used in an attempt to identify the different brain structures related to depression and to define *biomarkers*[10] for depression diagnosis.

5. https://www.verywellmind.com/structured-clinical-interview-2510532
6. https://www.geestelijkgezondvlaanderen.be/feiten-cijfers
7. https://www.rand.org/research/gun-policy/analysis/supplementary/mental-health-access-and-suicide.html
8. https://www.mayoclinic.org/diseases-conditions/multiple-sclerosis/diagnosis-treatment/drc-20350274
9. https://www.webmd.com/epilepsy/guide/electroencephalogram-eeg
10. https://dictionary.cambridge.org/dictionary/english/biomarker: Biomarker: something, for example a gene or substance, that shows that a particular biological process or condition is present.

EEG has a high temporal resolution and can be used to examine the possible changes in frequency and connectivity to develop a depression biomarker but has a low spatial resolution, making the identification of influences from specific brain regions difficult, if not impossible [23]. EEG based biomarkers that can diagnose people with depression with great accuracy have been developed; some results report accuracies ranging from 85% up to 99% [24], [25], [26], [27]. The absence of a train-validation split of the data set or cross-validation from some of these papers as well as the small data sets (30 to 60 people) that are used by all papers limit the value of these biomarkers.

Depression biomarkers based on fMRI data have also been developed. Both resting state and task-related fMRI data are used to define biomarkers. Resting state biomarkers should reflect the changes in resting state brain networks and could point to changes in self referential thinking/rumination (negative self-image) due to depression while task-related fMRI biomarkers reflect reactions to stimuli or tasks [28], [29]. Acuracies up to 92% have been reported for resting state fMRI based biomarkers [30], [31], [32]. Task-related fMRI biomarkers have been reported with accuracies up to 95%[33],[34] [35],[36].

Although resting state fMRI based depression biomarkers have a high accuracy, they all have two disadvantages that limit their potential use in practice. The first limitation is that all biomarkers have been built using a limited data set. This lowers the possibility that the biomarker detects changes that are due to depression and increases the chances of detecting imbalances in the data set instead. It should be noted that this is probably not the case, just that the possibility rises. A second limitation is that almost all biomarkers that have been developed are built using features that are not easily interpretable and thus provide little to no information about the underlying illness.

Both problems will be addressed in this master's dissertation. A large data set of 46 patients with depression and 60 healthy controls will be used to build a biomarker for depression. The features that are explored will be easily interpretable for a physician.

1.2 Objective

The objective of this master's dissertation is to build a computer-aided diagnosis tool capable of diagnosing depression based on an MRI scan of the brain, a resting state fMRI scan of the brain or a combination of both. The data set that is available contains 60 healthy controls and 46 patients diagnosed with depression. The master's dissertation is composed of nine chapters and consists of five main parts: a theoretical introduction, structural feature selection, functional feature selecton, classifier training and result interpretation.

Theoretical introduction

In the first part the basic principles of MRI and fMRI will be explained as well as some principles of machine learning, specifically support vector machines.

Structural feature selection

The second part is the selection of the structural features. These features reflect the structural aspects of the brain such as cortical thickness and volumes of different brain regions. These features are calculated from a T1 MRI scan using the FreeSurfer program [77].

Functional feature selection

The third part is the functional feature selection process. Two different feature types are explored: intensity features and connectivity features. Intensity features are features that incorporate the general activity in the brain, connectivity features reflect the functional connectivity variability in the brain.

A data-driven approach instead of a clinical approach is taken in the feature selection process. No prior knowledge about the pathology of depression is used to influence the feature selection process. The final used features are selected only due to their significance in the data set. Clinical relevance of the features, if any, is assessed after the features have been selected.

Classifier training

The fourth part is classifier training. A classifier training pipeline is defined and the feature sets obtained in the third part are used as input for the training of a classifier. Due to the limited size of the data set *support vector machines* are the main machine learning algorithm that is used as classifier.

Firstly training is done using a single feature type as input. This makes it possible to assess the viability of each feature type as a distinguishing factor in the classification task. The final result of this classification can give insight in possible underlying mechanisms of depression that may be previously not considered. General conclusions should be made very cautiously however, as the found results only reflect people in the used data set, not the general population.

At a later stage the most significant features of all feature types are combined and used to train the final classifier.

Result interpretation

The final part is the interpretation of the results found in part four. The accuracy of the different computed classifiers is compared and results are discussed.

Chapter 2

MRI and fMRI

In this chapter, the neuroimaging techniques that are used to measure the brain activity are explained. In the first part of the chapter the basic principles of MRI will be clarified while the second part of the chapter will describe the underlying principles of fMRI. This chapter is mainly based on the books *MRI from picture to proton* by DW McRobbie, EA Moore, MJ Graves and MR Prince [3] and *Functional magnetic resonance imaging* by SA Huettel, AW Song and G McCarth [39].

2.1 MRI

Magnetic resonance imaging, also known as MRI, is an anatomical imaging technique widely used in medicine to visualize the internal structures and organs of humans (and animals). The basic principle of MRI is *nuclear magnetic resonance* (NMR), which differentiates MRI from other medical imaging techniques such as X-ray radiography, computed tomography (CT), positron emission tomography (PET), single photon emission computed tomography (SPECT) and ultrasound.

2.1.1 Physical principles

2.1.1.1 Magnetic moment

Every atomic nucleus is built up from two subatomic particles: neutrons and protons. Both particles rotate around their own axis, creating a spin angular moment. This spin angular moment, S, is defined by a spin quantum number. As each particle can spin in only two possible directions, clock- and counterclockwise, the spin quantum number has only two possible values: $+\frac{1}{2}$ and $-\frac{1}{2}$.

Neutrons are electrically neutral, but protons do have a small positive electric charge. This electric charge can be modeled as a small current, I. The circuital law of Ampère states that a moving electrical current generates a magnetic dipole field B. This magnetic field is defined by its magnetic moment μ, given by formula 2.1.

$$\mu = \gamma.S \tag{2.1}$$

Here γ is the gyromagnetic ratio ($\frac{Hz}{T}$) and S is the spin angular moment defined by the spin quantum number. The magnetic moment can only occupy one of two possible states due to the limited possible values of the spin angular moment.

2.1.1.2 Total nuclear spin

The total nuclear spin of a nucleus depends on the total amount of protons and neutrons from which the nucleus is built up. When the nucleus consists of an even number of both neutrons and protons, the total nuclear spin is zero. When a nucleus consists of an even amount of one subatomic particle and an uneven amount of the other subatomic particle, the total nuclear spin will have a half integer value. When a nucleus consists of an odd amount of both subatomic particles the total nuclear spin will have a full integer value.

When the total nuclear spin of a nucleus is different from zero, the nucleus will have a magnetic moment. This magnetic moment is defined by formula 2.1, where the gyromagnetic ratio is nucleus specific.[11]

2.1.1.3 The larmor frequency

When a nucleus with a magnetic moment different from zero is objected to an external magnetic field (B), the nucleus will try to align with the magnetic field. Due to the fact that only two states are possible, the alignment of the nucleus with regard to the external magnetic field is not perfect: a small angle exists between the magnetic moment vector and the magnetic field vector.

When an angle exists between the magnetic field and the magnetic moment, the magnetic field exerts a torque on the magnetic moment of the nucleus, this torque (τ) is defined by formula 2.2.

$$\overrightarrow{\tau} = \overrightarrow{\mu} \mathrm{x} \overrightarrow{B} = \gamma . \overrightarrow{S} \mathrm{x} \overrightarrow{B} \tag{2.2}$$

Where $\overrightarrow{\tau}$ is the torque vector, $\overrightarrow{\mu}$ is the magnetic dipole moment, $\overrightarrow{B}$ the external magnetic field, γ the gyromagnetic ratio and $\overrightarrow{S}$ the angular momentum vector. This principle is shown in figure 2.1.[12]

Because of the torque, the angular momentum vector will start to precess around the external magnetic field vector. This precession occurs at a specific frequency, called the *larmor frequency*, given by formula 2.3.

$$\omega_L = \frac{1}{S.sin(\phi)} . \gamma . S . B . sin(\phi) = \gamma . B \ \ (rad/sec) \tag{2.3}$$

In this formula ϕ is the angle between B and μ (shown in figure 2.1).

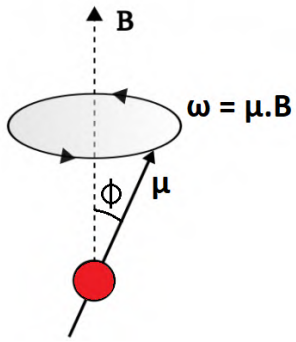


Figure 2.1: Visualization of the angle (ϕ) between the magnetic moment vector (μ) of a nucleus (shown in red) and an external magnetic field (B) and the corresponding precession.[10]

11. http://nmrwiki.org/wiki/index.php?title=Gyromagnetic_ratio
12. http://zerpoii.opentronix.com/?paged=4&tag=featured

Formula 2.3 shows that the larmor frequency depends on both the strength of the external magnetic field and the gyromagnetic ratio of the nucleus. The gyromagnetic ratio is unique for every nucleus type, so every nucleus type has a unique larmor frequency.

2.1.1.4 Relevance for MRI imaging

A single hydrogen atom is built up from a single proton and an electron. As described in section 2.1.1.3, this proton will have a small magnetic moment and will thus precess when an external magnetic field is present.

Hydrogen atoms account for roughly 62% of all atoms present in the human body [40]. When the human body is subjected to an external magnetic field, all hydrogen atoms in the body will start to precess. The gyromagnetic ratio of hydrogen is 42.58MHz/Tesla.[13] Note that not only hydrogen is subject to precession; other atoms such as carbon and nitrogen will also precess, but at another frequency [41]. When the precessing hydrogen atoms are excited (see section 2.1.2.2), a signal will be generated and can be measured. Different tissues in the human body contain different amounts of hydrogen atoms per volume. As each hydrogen atom generates a signal when excited, and more hydrogen atoms close together lead to a stronger signal, certain tissues generate stronger signals than others. Depending on the imaging sequence that is used to obtain the signal (see section 2.1.3), different signal strengths will translate into different grey values on the final MRI image. An example of an MRI image of a brain is given in figure 2.2.

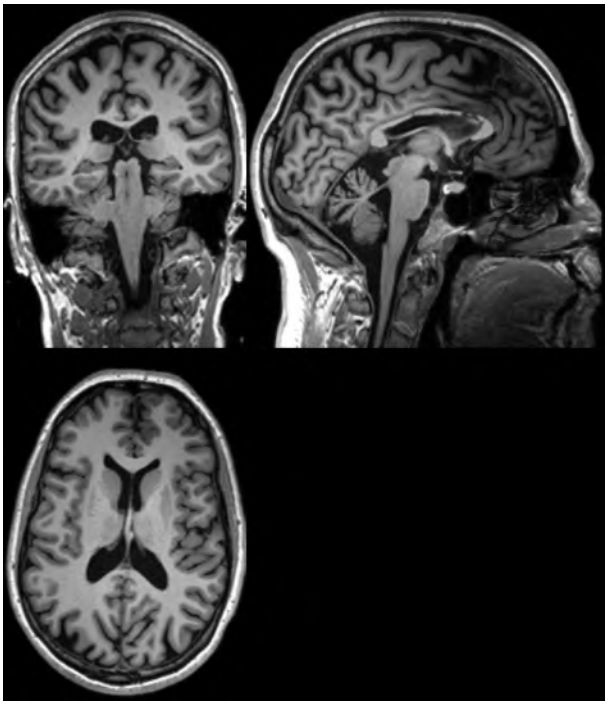


Figure 2.2: Example of an MRI image. The different tissues (grey matter, white matter, skin, air) can clearly be distinguished.

13. http://mriquestions.com/who-was-larmor.html

2.1.2 The MRI scanner

This section will explain the different parts of an MRI scanner and the influence they have on the generation of an MRI image. A simplified version of an MRI scanner is shown in figure 2.3.

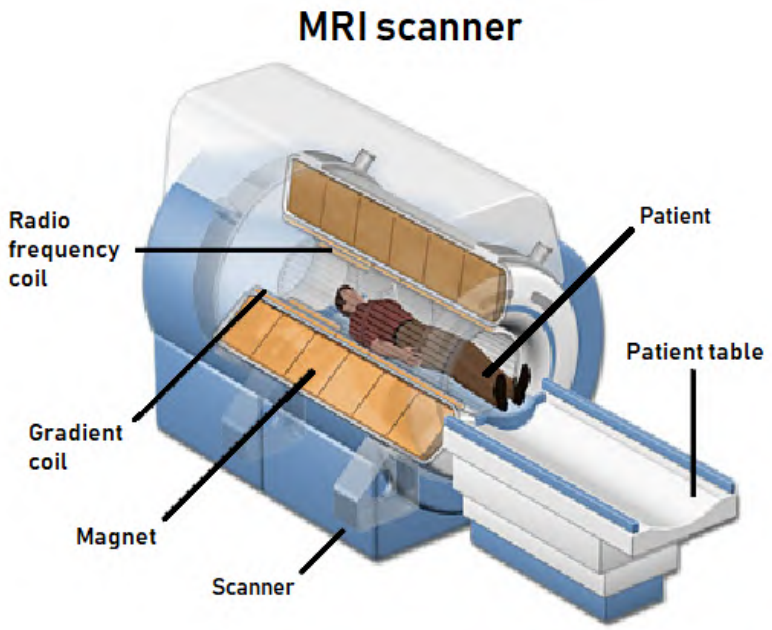


Figure 2.3: Simplified representation of an MRI scanner [2].

2.1.2.1 The homogeneous magnetic field

When no magnetic field is present, the magnetic moments of the hydrogen atoms in the body are orientated randomly, resulting in the absence of a net magnetization. The presence of a strong external magnetic field (B_0) results in the alignment of the magnetic moments to the magnetic field. Two possible alignments are possible: spin up and spin down. If a hydrogen atom has a spin up alignment the direction of the magnetic moment is parallel to the direction of the magnetic field, a spin down alignment means anti-parallel alignment. Because more hydrogen atoms are in a spin up state, a net magnetization (M) whose direction is parallel to B_0 is present. This principle is shown in figure 2.4.[14] The magnet that is used in an MRI scanner to generate B_0 is shown in figure 2.3 and is denoted as "Magnet". M is a static magnetic field and cannot be measured using a detection coil. Extra steps need to be taken in order to generate a measurable signal.

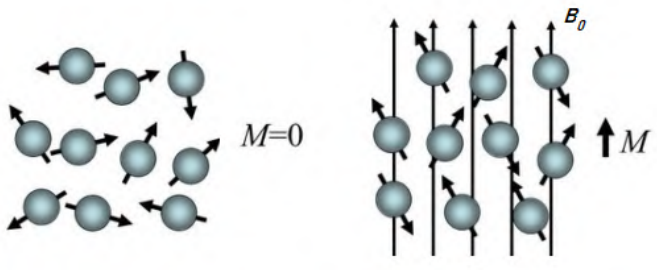


Figure 2.4: Random orientation of the hydrogen atoms without an external magnetic field (left), net magnetization M_L due to a magnetic field B_0 (right).

14. http://199.116.233.101/index.php/Physics_of_MRI

2.1.2.2 The rotating magnetic field

The net magnetization M cannot be measured if it stays static. To change the net magnetization a secondary magnetic field is used. The secondary magnetic field is a rotating field constructed by a combination of two radio frequency magnetic fields perpendicular to each other with changing field strength. The change in field strength follows a sine signal and both signals are offset by 90 degrees. The magnets generating this magnetic field are shown as "radio frequency coil" in figure 2.3. If the frequency of the rotating magnetic field is equal to the larmor frequency of a certain nucleus type, the net magnetization M of only that specific nucleus type will precess around B_0.

Due to the precession the net magnetization can be subdivided in two components: a longitudinal and a transverse component. The longitudinal component of the net magnetization (M_L) aligns with B_0 while the transverse component (M_T)is perpendicular to B_0. This is shown in figure 2.5. The angle between the magnetization and its longitudinal component, denoted as θ in figure 2.5, is called the flip angle. The flip angle is dependant on the duration of the radio frequency burst generated by the rotating magnetic field. A longer radio frequency burst results in a bigger flip angle.

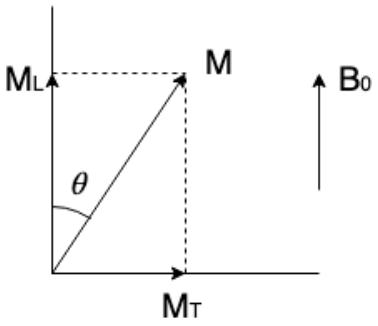


Figure 2.5: Visualization of the two components of the net magnetization M.

The longitudinal part of the net magnetization will remain stationary and can still not be measured, but the transverse part of the magnetization is measurable as it precesses at the larmor frequency and tries to align again with B_0. This signal is measurable because it induces an electric current. The part of the scanner that measures this current is called "scanner" in figure 2.3. The realignment of the transverse magnetization with B_0 is called relaxation. During the relaxation M_L will recover until it is completely restored (called longitudinal relaxation) and M_T will decay until none is left (called transverse relaxation). The relaxation of both magnetizations depends on different mechanisms that are independent of each other, resulting in different relaxation times that are shown in figure 2.6. The relaxation time of the longitudinal magnetization is called T1, the relaxation time of the transverse magnetization is called T2.

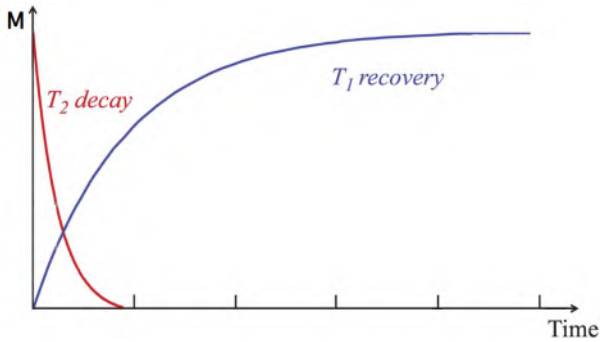


Figure 2.6: Visualization of the longitudinal (T1) and transverse (T2) recovery [3].

2.1.2.3 The gradient magnetic field

The homogeneous and rotating magnetic field makes it possible to generate a measurable signal, but all tissues in the body generate a signal at the same time and no comprehensible images can be generated. To solve this problem gradient magnetic fields are used; they are generated by gradient coils (see figure 2.3).

The gradient magnetic field generates a magnetic field that varies slightly in three dimensions. This magnetic field changes the local magnetic field strength of the homogeneous magnetic field slightly (see section 2.1.2.1). Each position of the body is now subjected to a slightly different magnetic field strength. As the larmor frequency depends on the gyromagnetic ratio of the nucleus, but also on the applied magnetic field strength (see formula 2.3), the hydrogen atoms at each postion of the body now have a slightly different larmor frequency. As nuclei only start to precess when the rotating magnetic field (see section 2.1.2.2) rotates at their larmor frequency, each position of the body can be targeted directly by changing the frequency of the rotating magnetic field. The position of the tissue that generates a signal can be encoded in the signal it generates. Now images can be created.

It should be noted that this is a simplified and incomplete explanation of the MRI scanner. For further details the reader is referred to the book *MRI from picture to proton* [3].

2.1.3 MRI imaging sequences

Each tissue type has different properties (such as amount of hydrogen atoms and the atoms in the lattice surrounding the hydrogen atoms) that influence the relaxation times. This difference in relaxation time results in differences in the measured signals. This difference is used to generate images. Many different imaging sequences that are able to visualize different aspects of the body by manipulating the T1 and T2 relaxations exist. These will not be explained here, but can be found in [3]. Only echo planar imaging (EPI) will be mentioned shortly as it is the imaging sequence that is used to obtain fMRI data.

2.1.3.1 Echo planar imaging

Echo planar imaging is an imaging sequence that is able to capture a complete 2D slice using a single radio frequency pulse generated by the rotating magnetic field. This is done by changing the gradient magnetic fields (see section 2.1.2.3) while the generated signal from the tissue is measured. This reduces the scanning time needed to measure the complete brain from tens of minutes to seconds. The main drawback of this method is the reduced resolution that is obtained. An example of a high resolution MRI image is shown in figure 2.2, an example of an EPI image is shown in figure 5.4. For more details the reader is referred to chapter 16, To BOLD go: new frontiers of *MRI from picture to proton* [3].

2.2 Functional MRI

Functional magnetic resonance imaging, also known as fMRI, is a functional imaging technique. Functional imaging is a form of imaging that does not reveal the anatomy of organs and structures, but reveals processes active in a person's body. The most common functional imaging techniques are functional MRI, positron emission tomography (PET), single photon emission computed tomography (SPECT), computed tomography perfusion imaging (CTPI) and near-infrared spectroscopy (NIRS). Functional imaging can provide information about illnesses that do not yet (or will never) alter the anatomy of a patient substantially enough that they could be detected using anatomical imaging. An example of this is the increased glucose uptake of small tumors in the body [42].

FMRI shows the activity of different brain regions both spatially and temporally with a normal MRI scanner (see section 2.1.2) using the EPI imaging sequence (see section 2.1.3.1). The underlying principle of fMRI is explained in section 2.2.1, the advantages and limitations of fMRI are discussed in section 2.2.2.

2.2.1 The BOLD response

FMRI shows the activity of the brain both spatially and temporally. It does this by measuring the differences in blood flow within the brain.

All cells in the body need energy to be able to perform functions. The creation of energy requires both oxygen and some energy source, mainly glucose. The citric acid cycle (also called Kreb cycle) will use the oxygen and energy source to produce adenosine triphosphate (ATP), the main energy source for the body. Neurons in the brain do not contain the necessary energy sources and oxygen themselves, so if they are active these resources are brought to them via the blood vessels in the brain through a process called the *hemodynamic response.* Local brain activity results thus in local variations of blood flow.

Oxygen is bound to haemoglobin in the blood and haemoglobin can exist in two possible states: oxygenated and deoxygenated. Oxygenated haemoglobin (called oxyhemoglobin) has different magnetic properties than deoxygenated haemoglobin (called deoxyhemoglobin) as oxyhemoglobin is diamagnetic and deoxyhemoglobin is paramagnetic. Blood traveling towards active neurons will contain a higher percentage of oxyhemoglobin relative to blood that delivered the oxygen to the active neurons. At the location of the active neurons oxyhemoglobin will become deoxyhemoglobin as oxygen is delivered to the neurons. This change of oxyhemoglobin/deoxyhemoglobin concentration shifts the magnetic properties of the blood from more paramagnetic to more diamagnetic. The magnetic property shift of the blood will be higher in locations where neurons are active compared to locations with inactive/less active neurons as more blood is delivered and more oxygen is given to the active neurons. This difference in magnetic properties is called the *blood-oxygen level dependant* (BOLD) response and is measurable using an MRI scanner.

The BOLD response changes through time and follows the hemodynamic response function closely. Its course through time is shown in figure 2.7.[15] Two observations can be made from this figure: the change in the BOLD signal is small (2%) when a stimulus is applied and takes a long time (25 seconds) to recover completely. Another observation is that the peak in the BOLD signal does not align in time with the stimulus but is delayed by around 8 seconds.

15. http://mriquestions.com/does-boldbrain-activity.html

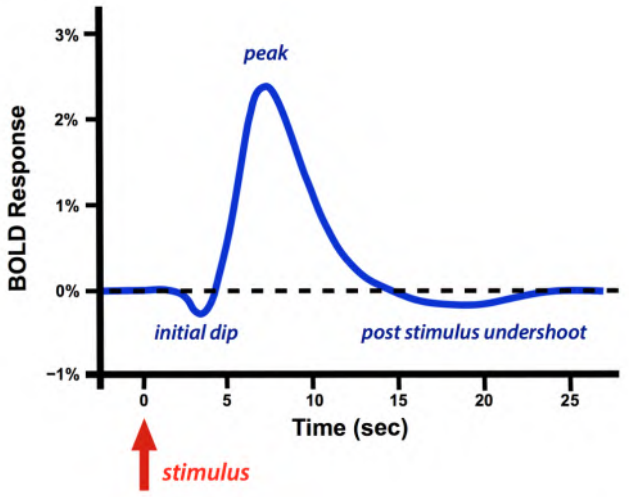


Figure 2.7: Visualization of the BOLD response with respect to time. The x-axis shows time, the y-axis the change in oxy-/deoxyhemoglobin concentration in the blood in percentage.

2.2.2 Advantages and limitations of fMRI

The main advantage of fMRI when compared to other neuroimaging techniques is its high spatial resolution of around 1-2mm. The high spatial resolution makes fMRI a unique imaging technique for identifying functional brain regions. Caution however should be used with the use of fMRI. It is not sure that the location where the oxygen is delivered in the brain corresponds perfectly with the location of the active neurons; the variations in the BOLD signal that are measured do not necessarily match perfectly with the associated brain regions. Another advantage of fMRI is its capability of recording brain activity of subcortical regions. Other (non-invasive) functional imaging techniques such as EEG and fNIRS are not capable of recording activity deep within the brain.

The major disadvantage of fMRI is its low temporal resolution when compared to other neuroimaging techniques (0.5-1Hz). Scanning the whole brain using an MRI scanner can take several minutes when an anatomical scan is needed[16] and the EPI imaging sequence (see section 2.1.3.1) that is used to obtain fMRI data significantly improves the temporal resolution, at the cost of spatial resolution. Other imaging techniques like EEG have a much higher temporal resolution (ms range). The low temporal resolution limits the research opportunities for fMRI. EEG and fMRI can however be recorded simultaneously and the data can be combined resulting in both high temporal and spatial resolution [39].

2.2.3 Resting state fMRI and task-related fMRI

Two types of fMRI data are defined: resting state and task-related fMRI data. Restig state fMRI is the capture of the BOLD signal changes when the person is not engaged in any activity. The person lies with his eyes closed in the MRI scanner and does not think about anything special. The person is not allowed to fall asleep. Several brain networks have been defined using resting state fMRI data such as the default mode network audio/visual networks and sensory/motor networks [39].

Task-related fMRI is the capture of the BOLD signal changes when a person performs certain tasks. The nature of these tasks is varied ranging from viewing images to counting tasks [35], [43]. This type of fMRI data is used to investigate the progression of signals through the brain and identify brain regions related to different tasks a person can perform.

16. https://www.healthline.com/health/head-mri

Chapter 3

Machine learning

This chapter will explain some principles of machine learning and the machine learning technique that is used: support vector machines (SVMs). Not all aspects of machine learning will be explained, only the principles that are of importance in this master's dissertation. This chapter is based on the book *Introduction to machine learning* by Ethem Alpaydin [44].

3.1 Introduction

Machine learning is the scientific study dedicated to the development and optimization of models and techniques used by computers to perform specific tasks without any specific instructions. This approach makes it possible to solve complex problems for which no obvious or simple instructions can be defined. The models instead try to learn underlying patterns using available data sets or by trial and error. Depending on the problem that needs to be solved, different models and approaches are required. The learning technique that is used in this master's dissertation is called supervised learning.

3.1.1 Supervised learning

This type of learning can be used when a mapping from input to output is needed and data that contains both the input and correct output, called labeled data, is available. The data set for this master's dissertation is labeled data. The input data is the MRI and fMRI scans, the output is the class they belong to: healthy controls or depression group. Many models can be trained with supervised learning such as linear regression, random forests, artificial neural networks and SVMs.

3.2 Support vector machines

An SVM is a machine learning technique that is a part of a group of machine learning models called kernel machines. SVMs are used in this master's dissertation because high classification accuracies can be obtained with small data sets (tens of samples) while other machine learning techniques, such as artificial neural networks, need much larger data sets (thousands of samples).

The general principle of SVMs will be explained by describing a simple classification problem. Two classes are present and two variables describe each data point belonging to one of the two classes. Class 1 will be denoted by C_1, class 2 by C_2. Data points belonging to C_1 are labeled with -1, data points belonging to C_2 with +1. A possible distribution of data points of both classes is shown in figure 3.1.

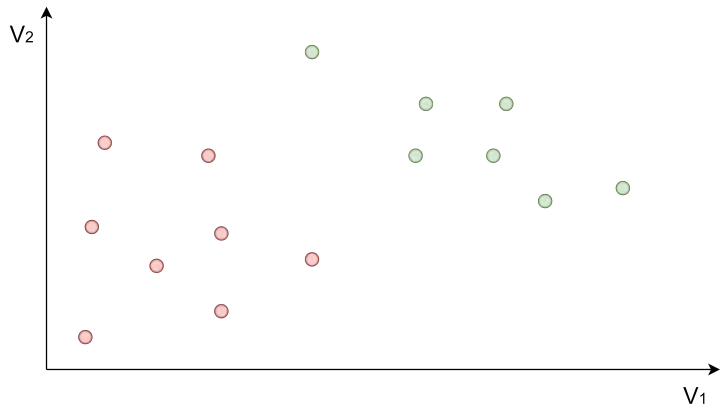


Figure 3.1: A possible distribution of data points. Data points beloning to C_1 are shown by red dots, C_2 is shown by green dots. V_1 and V_2 denote two variables.

The samples represented in figure 3.1 can be written as $X = \{\mathbf{x}^t, r^t\}$. X is a 2*t matrix containing every data point x coupled with its corresponding class label r, with t the amount of data points. The goal of a classification problem is to predict the class label r of a data point given the variables x that describe it. A solution for this problem is to define a hyperplane which separates both classes. As the current case is defined in two dimensions, the hyperplane will be a line.

The hyperplane can be defined by the function described in formula 3.1. It presents the hyperplane as a function of the available data points ($\mathbf{x}$), each multiplied by some weight ($\mathbf{w}$) where an offset is added (w_0).

$$g(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + w_0 \tag{3.1}$$

The hyperplane will be able to separate both classes when it obeys two constrains, shown in formula 3.2 and 3.3.

$$\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x}^{\boldsymbol{t}} + w_0 \geq +1 \text{ for } r^{\mathrm{t}} = +1 \tag{3.2}$$
$$\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x}^{\boldsymbol{t}} + w_0 \leq -1 \text{ for } r^{\mathrm{t}} = -1 \tag{3.3}$$

These constrains demand that for every data point belonging to C_1 the hyperplane will return a value ≥ 1 while returning a value $\leq$ -1 for every data point belonging to C_2. This equation can be simplified to formula 3.4.

$$r^{\mathrm{t}}(\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x}^{\boldsymbol{t}} + w_0) \geq +1 \tag{3.4}$$

It should be noted that this is a tough constraint to obey. Not only do we want a correct separation (which would require $r^{\mathrm{t}}(\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x}^{\boldsymbol{t}} + w_0) \geq 0$) but also that all points are some distance away from the hyperplane. The space between the defined hyperplane and the data points that are closest to it is called the margin. The best classification results will be obtained when the margin on both sides is maximized. If this is not possible, soft margin SVMs need to be used; these will not be explained but can be found in [44].

The maximization of the margins is obtained using formula 3.5 and describes the minimization of the norm of the weight vector $\boldsymbol{w}$. This minimization results in a description of the optimal hyperplane by the least amount of data points necessary as only the data points closest to the hyperplane will have a weight greater than zero. This results in a description of the optimal hyperplane by only a small subset of the initial present data points. These data points are called the *support vectors*. The optimization of

the weight vector is the "learning" part of this machine learning technique. A visual representation of the optimal hyperplane of the distribution shown in figure 3.1 is shown in figure 3.2. This example was a 2D classification problem as two variables were given. This principle can be extended to N variables.

$$min\left(\frac{1}{2}||\mathbf{w}||^2\right) \text{ subject to } r^{\mathrm{t}}(\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x}^{\boldsymbol{t}} + w_0) \geq +1 \tag{3.5}$$

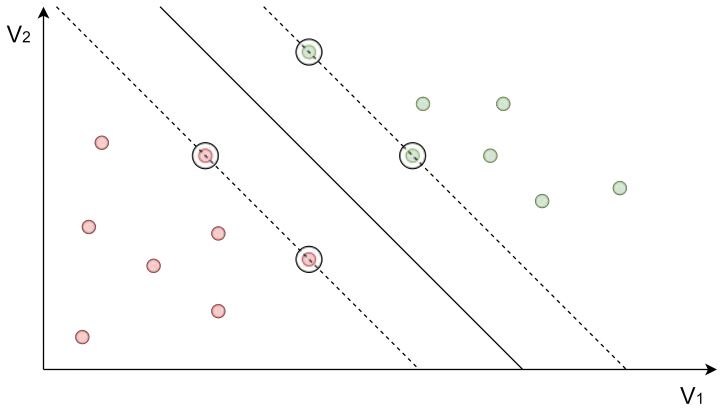


Figure 3.2: Visual presentation of the optimal hyperspace separating C_1 and C_2. The support vectors are encircled.

It should be noted that this explanation concerns the most simple case possible. More information about kernel models can be found in *Introduction to machine learning* [44].

Chapter 4

Structural features

This chapter describes the feature selection process that is used to obtain the structural features. First the specifics of the data set that is used are discussed. Secondly the program FreeSurfer [77], used to calculate the structural features, is described. Finally the clinical relevance of the found features is assessed.

4.1 Available dataset

The data used in this master's dissertation is obtained from Prof. Chris Baeken, one of the promotors.

The complete data set consists of two groups: healthy controls and patients with depression. The healthy control group consists of 60 individuals, the depression patients group consists of 46 individuals. The data set of the depressed patients consisted originally of fifty right-handed patients, of which 35 were female (average age 42 years, standard deviation (SD) = 12 years). All patients were antidepressant free at the time and were at least stage I treatment resistant, meaning that all of them had had at least one unsuccessful treatment trial with serotonin reuptake inhibitors or noradrenaline reuptake inhibitors. Further exclusion criteria were current or past history of epilepsy, neurosurgical intervention, having a pacemaker, having a metal object in the brain, having undergone electroconvulsion therapy, alcohol dependence or suicide attempts within 6 months before the start of study. Patients with co-morbidities such as bipolar disorder and psychosis were also excluded. Depression was diagnosed using the structured Mini-International Neuropsychiatric Interview. Four patients were not used; one female patient due to a suicide attempt (medication overdose), one female patient due to spontaneous improvement of the condition, one male patient due to an extra neurostimulation session and one patient (sex unknown) due to the absence of data. The healthy controls are matched for sex, age and education level. Further details regarding the data set can be found in [45].

The healty control group will be called *healthy controls*, the patients with depression the *depression group*.

Every person in the whole data set (106 people in total) has two different MRI scans: a T1 weighted gradient echo (GE) scan using the MPRAGE protocol [46] and an fMRI scan using the EPI imaging sequence.

4.2 Structural features

Structural features reflect possible anatomical changes due to depression and are obtained from the anatomical T1 images of the data set. The program that is used to obtain these features is called FreeSurfer [77].

Two feature subtypes will be explored: cortical thickness and parcel volume. Cortical thickness refers to the thickness of the gray matter (in mm) of certain parcels. Parcel volume refers to the total volume (in mm^3) of certain parcels. FreeSurfer uses a different brain atlas than the Lausanne brain atlas, so the structural features will not reflect properties of the same regions as the intensity and connectivity features. Less different regions are defined than in the Lausanne brain atlas: 35 cortical thickness parcels are defined for each hemisphere and 39 parcel volumes are defined. Due to the low amount of possible features compared to the other two feature types, no feature selection process will be used. Instead the significance of each possible feature will be tested by a two-pair t test. Any significant feature will be used in a feature set. Three feature sets are defined: a left cortical thickness feature set, a right cortical feature set and a parcel volume feature set. The clinical relevance of these features is discussed in section 4.2.2, the results of the feature sets are shown in section 7.3.1 and are discussed in section **??**. First the FreeSurfer program will be clarified.

4.2.1 FreeSurfer

A specific function from FreeSurfer is used, called *recon-all*. As it consists of a total of 29 steps[17], the steps will be clarified in a simple way and some steps will be put together as they are part of a larger step.

4.2.1.1 Step 1: Normalization

The first step in the FreeSurfer workflow is normalization. The anatomical MRI scan will be transformed into the MNI305 atlas space using an affine transformation. Intensity correction is also applied. A second intensity correction will be performed after step 4 as the exclusion of the skull improves intensity correction.

4.2.1.2 Step 2: Skull strip

The second step is called skull strip. As its name implies, the skull will be removed from the scan, leaving only the brain and neck.

4.2.1.3 Step 3: Subcortical segmentation

The third step is defining and segmenting the subcortical regions. This is done in multiple steps where the neck is stripped and several registrations to templates are made. This step ends with segmented and labeled subcortical regions.

4.2.1.4 Step 4: Statistics calculation

The fourth step is the calculation of the statistics of the subcortical parcels. The volume of these subcortical parcels, used as parcel volume features, are calculated in this step.

17. https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all

4.2.1.5 Step 5: White matter segmentation

The fith step is the segmentation of the white matter. The constraints used for this segmentation are the intensity difference between gray and white matter, the immediate neighboring voxels around each voxel and the inherent smoothness of the border between white and grey matter.

4.2.1.6 Step 6: Brain division

The sixth step is the division of the brain in the left and right hemisphere, the cerebellum and the midbrain.

4.2.1.7 Step 7: Tesselation

The seventh step is the tesselation of both hemispheres. The surface of each hemisphere is approximated using a finite element method where the border, defined in step 5, is approximated using small triangles. The smallest edges of the triangles have the same length as the side of a voxel. A surface is defined this way.

4.2.1.8 Step 8: Surface smoothing

The eight step is the smoothing of the surface defined in step 7. As the surface follows the voxels faces which define the surface the edges will be perpendicular to each other. Smoothing the edges reduces the angle and makes the surface smoother.

4.2.1.9 Step 9: Inflation

The ninth step is the inflation of the smoothed surface. The surface will be inflated to smoothen the gyri and sulci and the transformation for each vertex and edge is calculated. When the inflated surface is obtained, it will be checked for any defects that are present due to errors in previous steps.

4.2.1.10 Step 10: Surface definitions

The cortical thicknesses are defined by aligning the inflated surface with the grey-white matter border present in the original anatomical T1 MRI scan. A second surface that defines the pial surface is created.

4.2.1.11 Step 11: Spherical inflation

The inflated surface is further inflated until it becomes a sphere. This spherical surface is matched to a spherical atlas defining the different brain regions. Alignment is performed based on matching folding patterns of the brain and afterwards using small scale patterns.

4.2.1.12 Step 12: Parcel labeling and statistics calculation

The different parcels will be labeled and the statistics of of the parcel will be calculated. The cortical thicknesses of these parcels, used ase left and right hemisphere thickness features, are calculated in this step.

4.2.2 Interpretation of the found features

A total of three feature sets (left hemisphere thickness, right hemisphere thickness and parcel volume) have been defined and will be discussed. The results of the classification process can be found in section 7.3.1.

Interpretation of the used tables

The *number* column defines the significance of the feature (here the features with the lowest p-value when a double-pair t test is performed). The feature with number 1 has the lowest p-value of all features. The feature with number 10 had the tenth lowest p-value of all features. The *parcel* column defines the name of the parcel whose thickness or volume is linked with the feature. The *sign* column defines which group (healthy controls or depression group) has the biggest cortical thickness or parcel volume. A "+" sign denotes that the healthy controls have on average a bigger cortical thickness or parcel volume, a "-" sign denotes that the depression group has on average a bigger cortical thickness or parcel volume. It should be noted that this does not occur.

4.2.2.1 Left hemisphere thickness

Only nineteen features are found to be statistically significant, they are shown in table 4.1. Reduced cortical thickness of multiple brain regions defined by the features have been reported in literature. The brain regions from which a reduced cortical thickness is closely linked with depression (as indicated by prof. Baeken) are the rostral middle frontal gyrus (feature 2), the precentral gyrus (feature 6), the insula (feature 7), the precuneus (feature 10), the pars orbitalis (feature 14), the frontal pole (feature 16), the superior frontal gyrus (feature 17), the post central gyrus (feature 18) and the caudal middle frontal gyrus (feature 19) [78], [79], [80]. The presence of several brain regions associated with depression shows the clinical relevance of the left hemisphere thickness features. The results of the classifiers trained with the left hemisphere features are shown in section 7.3.1.1 and are discussed in section **??**.

Table 4.1: Significant left hemisphere thickness features.

Number	Parcel	Sign
1	Pars opercularis	+
2	Rostral middle frontal gyrus	+
3	Superior temporal gyrus	+
4	Mean left hemisphere thickness	+
5	Supramarginal gyrus	+
6	Precentral gyrus	+
7	Insula	+
8	Pars triangularis	+
9	Inferior temporal gyrus	+
10	Precuneus	+
11	Inferior parietal gyrus	+
12	Middle temporal gyrus	+
13	Lateral orbitofrontal cortex	+
14	Pars orbitalis	+
15	Fusiform	+
16	Frontal pole	+
17	Superior frontal gyrus	+
18	Post central gyrus	+
19	Caudal middle frontal gyrus	+

4.2.2.2 Right hemisphere thickness

Only eleven features are found to be statistically significant, they are shown in table 4.2. Only two of them can consistently be linked with depression (as indicated by prof. Baeken): the frontal pole (feature 5) and the superior frontal gyrus (feature 10) [79], [80]. From a clinical relevance point of view, the right hemisphere features are thus less significant than those of the left hemisphere. This observation, together with the fact that less statistically significant features have been found in the right hemisphere could show that a difference exists between the left and right hemisphere. An imbalance between a hypoactive left and a hyperactive right hemisphere in people with depression has been reported multiple times and these observations could also point towards this imbalance [52], [53].

Table 4.2: Significant right hemisphere thickness features.

Number	Parcel	Sign
1	Supramarginal gyrus	+
2	Inferior parietal gyrus	+
3	Pars triangularis	+
4	Mean right hemisphere thickness	+
5	Frontal pole	+
6	Inferior temporal gyrus	+
7	Superior temporal gyrus	+
8	Superior parietal gyrus	+
9	Pars opecularis	+
10	Superior frontal gyrus	+
11	Middle temporal gyrus	+

4.2.2.3 Parcel volume

Only six features were found to be statistically significant, they are shown in table 4.3. The first two features are both hemispheres of the cerebellum, indicating that the biggest change in volume between the healthy controls and the depression group could be a decrease in volume in the cerebellum for people with depression. Reduced cerebellar volume and cerebellar atrophy in people with depression is described in literature and an involvement of the cerebellum in several psychiatric disorders including depression is suspected and investigated [81], [82], [83]. The third and fourth features are both parts of the caudate nucleus. The involvement of the caudate nucleus in depression has been proposed as some diseases involving the caudate nucleus, such as caudate infarcts or Huntington's disease, give rise to depressive symptoms. Reduced volume of both caudate nuclei has been reported in people with depression [84], [85]. Reduced caudate volume has also been reported in other psychiatric disorders such as obssesive compulsive disorder [86]. Reduced thalamic volume is also linked with depression [87]. From a clinical relevance point of view, the parcel volume features are highly significant. The results of the classifiers trained with the parcel volume features are shown in section 7.3.1.3 and discussed in section **??**. The sixth feature is closely linked with the first two features.

Table 4.3: Significant parcel volume features.

Number	Parcel	Sign
1	Right cerebellum cortex	+
2	Left cerebellum cortex	+
3	Right caudate nucleus	+
4	Left caudate nucleus	+
5	Right thalamus	+
6	Right cerebellum white matter	+

Chapter 5

Functional features

This chapter will describe the calculation and clinical validation of the functional features. Two different functional features are calculated: intensity and connectivity features. Intensity features reflect the activity of the brain, connectivity features reflect the functional connections within the brain. Firstly the preprocessing process will be described, secondly the feature selection process is explained, thirdly the clinical validation of the obtained functional features is discussed.

5.1 Preprocessing process

This section will describe all preprocessing steps taken to prepare the data set for feature selection. The process is shown in figure 5.1. Several preprocessing steps can introduce errors in the data. To counter this, visual quality checks are performed. This is done by viewing the produced results of each preprocessing step of around ten people from each group.

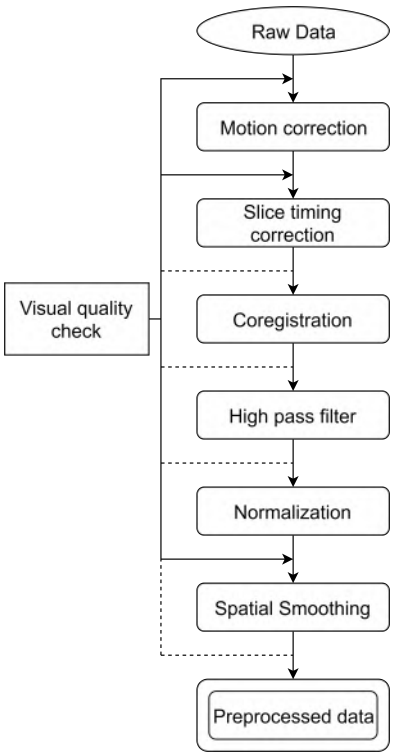


Figure 5.1: The complete preprocessing pipeline.

5.1.1 Step 0: File conversion

The MRI scanner used to obtain the scans is a Siemens Magnetom TrioTim syngo MR B17 scanner. The file format of the scanner is the *digital imaging and communications in medicine* (DICOM) with file extension .IMA. The zeroth preprocessing step is the conversion from the Siemens DICOM format to the *neuroimaging informatics technology initiative* (NIfTI) format. This step is done by the SPM toolbox ([37]) for further processing as the toolbox uses the nifti format. This step is performed using the *import.dicom* function from the SPM toolbox.[18]

5.1.2 Step 1: Removal of the first and last scans for signal equilibrium

The first five scans from each patient are dismissed, this because the BOLD signal (see section 2.2.1) has not yet been stabilized due to an incomplete T1 relaxation (see section 2.1.2.2 and figure 2.6). These scans are called dummy scans. This principle is shown in figure 5.2.

The last five scans are also dismissed, this because the final scans of some patients show a significant drop in voxel value throughout the whole scan. The origin of this signal drop is unknown.

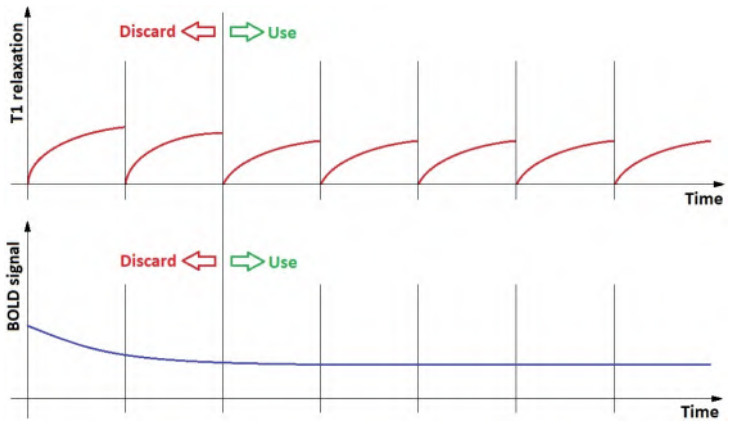


Figure 5.2: Visual representation of the T1 relaxation of the tissue and the corresponding BOLD signal. The vertical lines show the start of a new scan. (based on [4])

5.1.3 Step 2: Motion correction

Even though people are instructed to lie still when a scan is taken, small movements are inevitable as movements due to breathing and cardiac pulse also influence the location of the patient within the scanner. To correct these small movements, motion correction is applied.

Motion correction is done using a six parameter rigid body transformation. The first three parameters describe the rotation within the 3D space, the last three the translation. The appropriate rotation and translation values, which are unique for each scan, are calculated using a least squares approach [47]. This preprocessing step is performed using the SPM *realign* function.[19] This preprocessing step and all folowing steps aside from the high pass filtering (see section 5.1.6), are performed using the CONN toolbox [38].

18. https://en.wikibooks.org/wiki/SPM/Importing_data_from_the_scanner
19. https://en.wikibooks.org/wiki/Neuroimaging_Data_Processing/Realignment

5.1.4 Step 3: Slice timing correction

The next step in the preprocessing process is slice timing correction. This step is needed because it is assumed that a single EPI (see section 2.1.3.1) scan of the head is done instantaneously. In practice however, this is not correct. The EPI sequence used to obtain the fMRI data, scans the head in a sequence of several 2D slices. The acquisition of the 2D slices is not simultaneous but is done sequentially. This results in a single 3D scan where the values of each slice are obtained at a slightly different time.[20]

The solution to this problem is the interpolation of the measured values to correct for the time differences between the acquisitions, called slice timing correction. The time differences are defined by the repitition time (TR) and acquisition time (TA) parameters of the EPI sequence. The TR parameter defines the time between radio frequency bursts generated by the rotating magnetic field (see section 2.1.2.2). The TA parameter defines the time difference between the start of the first and last slice acquisition and can be calculated by formula 5.1.[21] The values for TR can be found in appendix A.

$$TA = TR - (\frac{TR}{nSlices}) \tag{5.1}$$

The last two parameters that are needed are the reference slice and the slice order. The reference slice is the slice that is firstly scanned and is dependent on the slice order ie. the order in which the different slices are acquired. The used slice order is an *interleaved* slice order (see *"Series"* in appendix A), meaning that first all even or uneven slices are acquired after which the other half are captured. As the scanner used to obtain the fMRI data is a *Siemens Magnetom TrioTim syngo MR B17* (see appendix A), the slice order is dependent on the amount of slices. If the total number of slices is even, the even slices are captured first; if the total number of slices is uneven, the uneven slices are captured first.[22] The amount of slices is 40 (for this data set, see appendix A) so the even slices will be captured first. This also defines the reference slice as slice two. This part of the preprocessing part is done using the *temporal.st* function from SPM, built into the preprocessing pipeline of the CONN toolbox [38].

5.1.5 Step 4: Coregistration

The next step in the preprocessing is coregistration. Coregistration is the alignment of the functional data (fMRI data) to the structural data (or vice versa) so that they share the same coordinate space. This means that the time series in the fMRI data are now linked spatially with their corresponding brain region in the structural MRI. Coregistration is a patient-specific preprocessing step where the functional data of each patient is mapped to their personal T1 MRI scan. Coregistration is performed using a six parameter rigid body transformation (similar to motion correction, see section 5.1.3) where three parameters describe the applied translation and the other three the applied rotation. The *coregister* fuction from the SPM toolbox is used, which is built into the CONN toolbox [38].

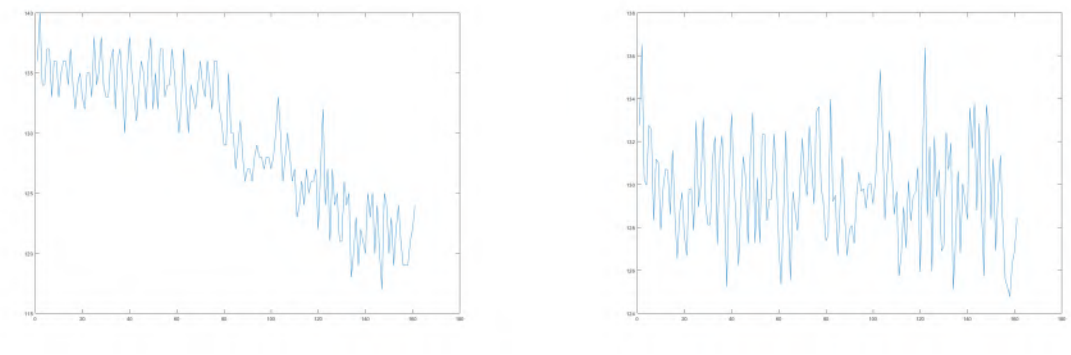
5.1.6 Step 5: High pass filtering

The next step is high pass filtering and is done to eliminate scanner drift. Scanner drift is the introduction of a low frequency signal (0 - 0.01Hz) into the time series captured during an MRI scan. The origin of scanner drift is a change in resonant frequency (see section 2.1.1.3) of the hydrogen protons

20. https://en.wikibooks.org/wiki/SPM/Slice_Timing
21. https://andysbrainblog.blogspot.com/2012/11/slice-timing-correction-in-spm.html
22. https://www.siemens-healthineers.com/magnetic-resonance-imaging/magnetom-world/clinical-corner/application-tips/slice-order-fmri.html

which is induced by changes in field strength of the static magnetic field (see section 2.1.2.1) [39]. Even though scanner drift is the main source of low frequency noise, other sources also influence the acquired signal [48].

The filter that is implemented has a bandstop frequency of $\frac{1}{128}$Hz and a bandpass frequency of $\frac{1}{120}$Hz. The presence of scanner drift and the result of the filtering is shown in figure 5.3. The function *myfreqfilter*[23] is used for the implementation of the high pass filter.



(a) Original time series. **(b)** Filtered time series.

Figure 5.3: Visualization of the initial presence of scanner drift (a) and the result after high pass filtering (b).

5.1.7 Step 6: Normalization

Normalization is the second registration step (after coregistration) in the preprocessing process. Again a body transform will be applied to the functional data of each person so that all people in the data set share a common coordinate space. This is necessary to allow comparisons between patients and find common characteristics on a group level. The coordinate space to which the data will be registered is the *Montreal Neurological Institiute* (MNI) space [49]. This step is performed by wrapping the personal T1 MRI scan of each patient to a the MNI template scan. Secondly the calculated transformation from this wrapping is applied to the functional data of each person.

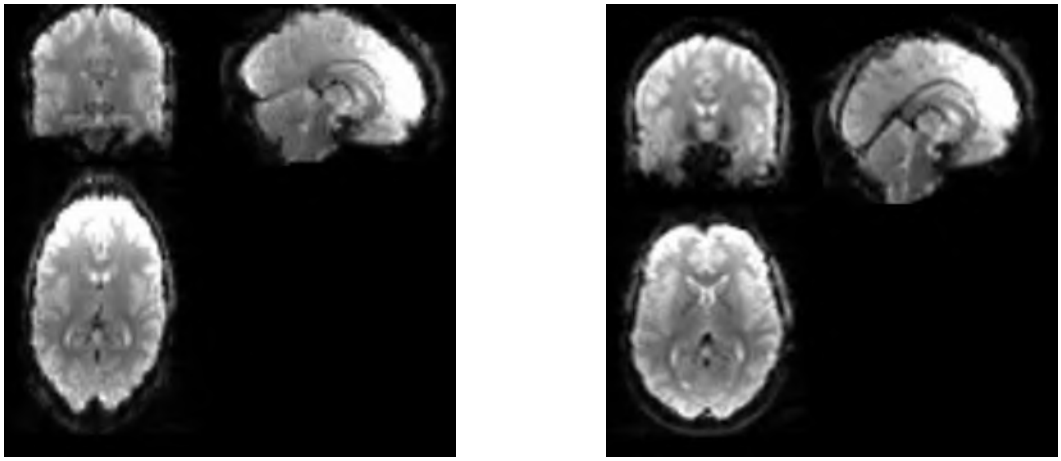
A first step in the normalization process is segmentation, where personal probability maps for air, skull, cerebrospinal fluid (CSF), white matter and gray matter are made. These probability maps are calculated from the T1 MRI scans of each person.

The second step is linear registration. The personal probability maps are aligned to template probability maps using an affine transformation, the template probability maps are located in the MNI coordinate space. The affine transformation has twelve parameters: three rotation parameters, three translation parameters, three zoom parameters and three shear parameters. The rotation and translation parameters align the personal probability maps to the template probability maps; the zoom and shear parameters alter the shape of the personal probability maps so that they are as similar as possible to the template probability maps.

The third and final step of the normalization process is non-linear registration. This registration is necessary as not all person-specific variations, such as unique folding of gyri and sulci, can be corrected using affine transformations. The non-linear registration is performed using a linear combination of

23. "myfreqfilter", Anthony Barone, The University of Texas at Austin, Institute for Geophysics.

discrete cosine basis functions. The result of this step is shown in figure 5.4. This step is performed using the *segment* and *normalize* function from SPM, which is used by the CONN toolbox [38], [37].

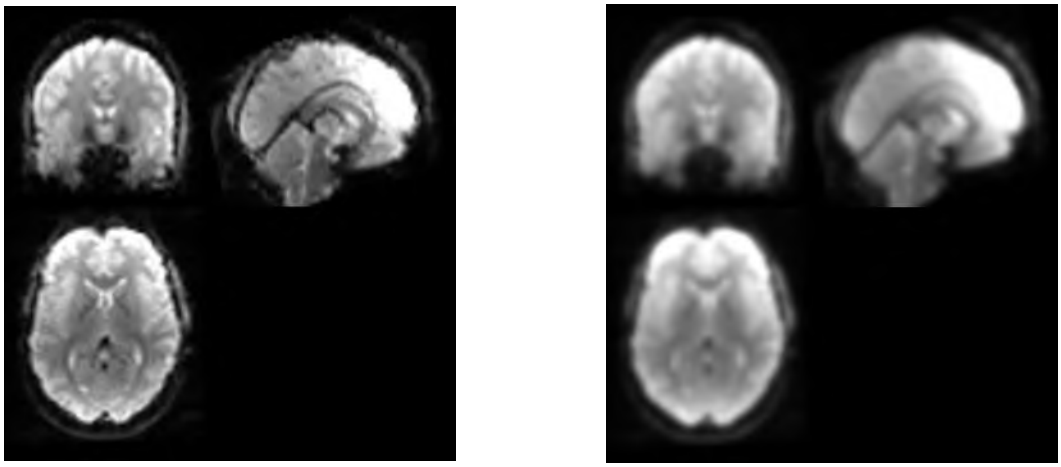


(a) Functional data before normalization. **(b)** Functional data after normalization.

Figure 5.4: Visualization of the effect of normalization.

5.1.8 Step 7: Spatial smoothing

The final step in the preprocessing process is spatial filtering, also called smoothing. This step is performed to increase the signal-to-noise ratio. Noise is present in fMRI data, but follows a (mostly) Gaussian distribution with an average value of zero.[24] As the signal due to neuronal activity is on average non-zero, the functional data will be spatially filtered using a Gaussian kernel. The kernel width is chosen to be 6 mm, a common choice in literature [30], [31]. The results of the spatial smoothing are shown in figure 5.5. This step is performed using the *smooth* function from SPM, which is used by the CONN toolbox [38], [37].



(a) Functional data before spatial smoothing. **(b)** Functional data after spatial smoothing.

Figure 5.5: Visualization of the effect of spatial smoothing.

24. http://mindhive.mit.edu/node/112

5.2 Parcellation

5.2.1 Human brain atlas

Both the intensity features and the connectivity features use a human brain atlas. A human brain atlas is a representation of the human brain where the brain is subdivided in several brain regions, also called parcels. The Lausanne brain atlas is used [50]. The Lausanne brain atlas is defined in the MNI space (see section 5.1.7) and has five different levels of brain region division. The first level divides the whole brain in 89 parcels, the second level has a division of 129 parcels, the third level has a division of 234 parcels, the fourth level has a division of 463 levels and the fifth level has a division of 1015 parcels.

Both the intensity and connectivity features will be calculated for each atlas level separately. This is done to investigate the influence of parcel size on the resulting features.

5.2.2 Human brain atlas resizing

The preparatory step that is taken is the resizing of the Lausanne brain atlas. The Lausanne brain atlas has been built using structural MRI images and has a higher resolution than the fMRI data from the data set. All atlas levels will be resized to match the fMRI data using the *imresize3* function from MATLAB.[25] The nearest neighbor option for interpolation is used.

5.3 Intensity features

The first category of features that will be explored are intensity features. They reflect the average activity of the different brain regions during resting state and are obtained from the fMRI data. It should be noted that these features only give a simplified representation of the brain activity as not all information in the fMRI data is used.

5.3.1 Absolute intensity and relative intensity

Two intensity feature subtypes will be explored: absolute and relative intensity features. The absolute intensity features are the features that reflect the average intensity of each brain region. They are found by averaging all time series to a single value for each time series and then averaging all values of a single brain region that is defined by the Lausanne brain atlas. It should be noted that this type of feature is sensitive to differences between patient scans. A global elevation or decrease of the recorded values within the scan of a patient will influence the final feature values of the patient significantly. Relative intensity features counter this problem.

Relative intensity features also reflect the average activity of brain regions through time. The main difference with absolute intensity features is the fact that the intensity of the different brain regions is normalized for each patient personally. This is done using formula 5.2. Here $I_{relative,j}$ is the relative intensity feature value of brain region j, $I_{absolute,j}$ is the absolute intensity feature value of brain region j and $I_{brain,average}$ is the average intensity value of the whole brain. $Ibrain, average$ is calculated by averaging all values of all time series.

$$I_{relative,j} = \frac{I_{absolute,j} - I_{brain,average}}{I_{brain,average}} \tag{5.2}$$

25. https://www.mathworks.com/help/images/ref/imresize3.html

5.3.2 Feature selection process

The feature selection process is shown in figure 5.6. The process starts with the preprocessed data (see section 5.1) and can be divided in two parts: a patient specific and a group specific part. The patient specific part defines the steps that are done on the data of each patient separately, the group specific part defines the steps that are done on a group level (healthy controls and depression patients). Each different step will be explained.

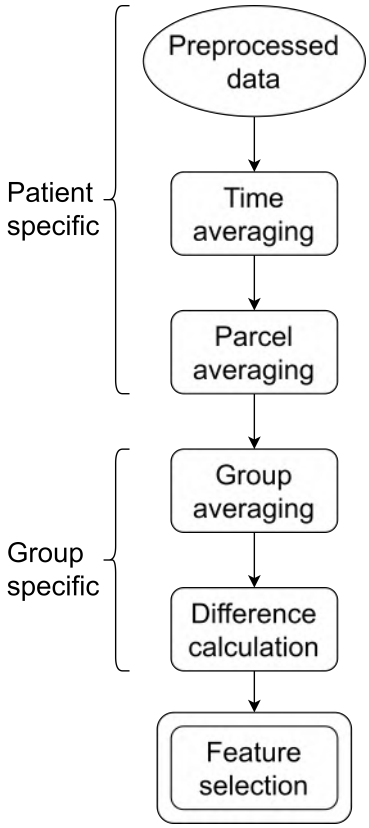


Figure 5.6: The intensity feature selection process.

5.3.2.1 Time averaging

The first step of the patient specific part of the feature selection process is time averaging. The time series of the fMRI data of each patient will be averaged. This reduces the patient data matrix from a 30076x160 or 30076x290 matrix (30076 equals the amount of time series that overlap with the Lausanne brain atlas, 160 or 290 equals the amount of different time points in an fMRI time series) to a 30076x1 matrix.

5.3.2.2 Parcel averaging

The second step of the patient specific part is parcel averaging. The 3D location of each time series (now average value) is mapped to the Lausanne brain atlas and linked to the corresponding parcel. The different time series that belong to the same parcel are averaged, resulting in a single average intensity value for each parcel. This reduces the patient data matrix from a 30076x1 matrix to a Nx1 matrix (N equals the amount of parcels of each atlas level, see section 5.2.1). As five atlas levels exist and both absolute and relative intensity values are calculated, ten different data matrices are calculated for each patient.

5.3.2.3 Group averaging

The third step of the feature selection process and first step of the group specific part is group averaging. The value of each parcel is averaged for both groups (healthy controls and depression patients). This results in 10 Nx1 matrices for each group.

5.3.2.4 Difference calculation

The fourth step of the feature selection process is difference calculation. The average group values of the depression patients are subtracted from the average group values of the healthy controls. This results in ten Nx1 matrices called difference matrices.

5.3.2.5 Feature selection

The final step is the feature selection itself. The 20 parcels that have the biggest absolute value in the difference matrices are selected. These parcels have the biggest different intensity values on average between the healthy controls and depression patients and are therefore best suited to be used as features to distinguish both groups. The significance of the features was validated statistically using the two-sample t-test, done by the *ttest2* function from MATLAB.[26]

As five atlas levels are used and two feature subtypes are calculated, ten feature sets are obtained. Each feature set is a 106x21 matrix; 106 equals the total amount of people in the data set and 21 equals the twenty features plus a final value defining the group to which the person belongs. All feature sets are normalized using the z-score.

26. https://www.mathworks.com/help/stats/ttest2.html

5.3.3 Interpretation of the found features

As five atlas levels are used and two feature subtypes are calculated, ten feature sets are obtained. Two sets will be discussed in this section: the best performing feature set of each intensity feature subtype. The results of the classification process using these two feature sets can be found in section 7.3.1, the discussion of the results in section ??.

5.3.3.1 Interpretation of the used tables

The *Number* column defines the significance of the feature. Feature 1 defines the parcel that has the biggest value in the difference matrix, feature 20 defines the parcel that has the twentieth biggest difference value in the difference matrix (see section 5.3.2.4). The *Parcel* column defines the name of the parcels that are related to the features. These are the parcels whose feature values differ on average the most between the healthy controls and the depression group. "(L)" denotes that the parcel is located in the left hemisphere, "(R)" denotes that the parcel is located in the right hemisphere. The number at the end of the parcel name is its further subdivision number (brain regions are subdivided in smaller regions when higher atlas levels are used). The *Sign* column defines the sign of the difference value. A "+" sign denotes a positive difference value, meaning that the intensity value of the parcel and by extension the general activity of that parcel is higher in the healthy controls compared to the depression group. A "-" sign defines the contrary.

5.3.3.2 Absolute intensity features

The twenty most significant features are shown in table 5.1. A first noticeable fact is that several features are subdivisions of a single brain region. Five features (feature 1, 9, 12, 18 and 20) are a subdivision of the left superior frontal gyrus (LSFG), three features (feature 3, 6 and 11) are a subdivision of the right superior frontal gyrus (RSFG), two features (feature 4 and 15) are a subdivision of the right rostral middle frontal gyrus (RRMFG) and six features (feature 5, 7, 8, 10, 13 and 14) are a subdivision of the left rostral middle frontal gyrus (LRMFG). It should be noted that all but one of these features (feature 7) are denoted by a "+" sign, meaning that all but one of these regions show hypoactivity in the depression group compared to the healthy controls. This shows that hypoactivity in the frontal lobe is clearly present in both hemispheres in this data set. This phenomenon has been reported in literature [51]. A second observation that is made is that the four brain regions that contain multiple features are two corresponding brain regions lying in both hemispheres, showing again the strong localization of (significant) hypoactivity to the frontal regions in the depression group when compared to the healthy controls. The difference in feature amount between these regions with respect to the hemisphere they are located in could point to the frontal asymmetry commonly found in people with depression [52], [53]. This difference in feature amount between the left and right hemisphere is also observed with the structural features (see section 4.2.2.2).

All four brain regions (LSFG, RSFG, LRMFG, RRMFG) are located in the prefrontal cortex. Hypoactivity in the prefrontal cortex is common in people with depression and has been described multiple times [54], [55], [56]. The LSFG specifically is a very significant brain region related to depression as it is one of the main regions for *transcranial magnetic stimulation* (TMS). The first (and therefore most significant) feature that is defined is the third subdivision of the LSFG and is located at the border between Brodmann area 10 and 46; this location almost exactly matches one of the locations for TMS for patients with depression [57], [58].

Another observation that can be made is the fact that both frontal poles are features that, contrary to almost all features, have a negative sign. This means that both frontal poles show increased activity in the depression group when compared to the healthy controls. This seems counter-intuitive as the

prefrontal cortex, which contains the frontal poles, is hypoactive in people with depression [54]. No clear explanation for this is currently known.

It can be concluded that, from a clinical relevance point of view, the absolute intensity features are highly significant.

Table 5.1: Absolute intensity features (Atlas3).

Number	Parcel	Sign
1	Superior frontal gyrus 3 (L)	+
2	Frontal pole (L)	-
3	Superior frontal gyrus 3 (R)	+
4	Rostral middle frontal gyrus 1 (R)	+
5	Rostral middle frontal gyrus 4 (L)	+
6	Superior frontal gyrus 4 (R)	+
7	Rostral middle frontal gyrus 6 (L)	-
8	Rostral middle frontal gyrus 2 (L)	+
9	Superior frontal gyrus 5 (L)	+
10	Rostral middle frontal gyrus 3 (L)	+
11	Superior frontal gyrus 2 (R)	+
12	Superior frontal gyrus 2 (L)	+
13	Rostral middle frontal gyrus 1 (L)	+
14	Rostral middle frontal gyrus 5 (L)	+
15	Rostral middle frontal gyrus 3 (R)	+
16	Frontal pole (R)	-
17	Caudal middle frontal gyrus (L)	+
18	Superior frontal gyrus 4 (L)	+
19	Precentral gyrus (R)	+
20	Superior frontal gyrus 9 (L)	+

5.3.3.3 Relative intensity features

The twenty most significant features are shown in table 5.2. A similar observation as with the absolute intensity features (see section 5.3.3.2) is made; most features are subdivisions from a few larger brain regions. The larger brain regions are the same as the brain regions found with the absolute intensity features. Seven features (feature 1, 2, 4, 6, 9, 11 and 15) are a subdivision from the left superior frontal gyrus (LSFG), four features (feature 7, 10, 12 and 20) are a subdivision from the left rostral middle frontal gyrus (LRMFG), three features (feature 3, 8 and 13) are a subdivision of the right superior frontal gyrus (RSFG) and two features (feature 14 and 19) are a subdivision of the right rostral middle frontal gyrus (RRMFG). The clinical significance of these features (as well as the left frontal pole, feature 5) has been discussed in the previous section (section 5.3.3.2) and will not be repeated.

The absolute intensity and relative intensity features are largely located in the same brain regions. This is expected, but strengthens the validity of the absolute intensity features. While the relative intensity features reflect the relative activity of a single brain region with respect to the global average activity, absolute intensity features only reflect absolute activities (see section 5.3.1). The similar features show that the absolute intensity features do not suffer from the lack of a normalization. It should be noted that this conclusion is only valid for the used data set. Any extension of this feature type to other data sets should be validated with a similar absolute versus relative intensity comparison. A difference between the absolute and relative intensity features is the presence of two features (feature 14 and 19), located in the RRMFG, that have a negative sign. No clear explanation for this is currently known.

Table 5.2: Relative intensity features (Atlas5).

Number	Parcel	Sign
1	Superior frontal gyrus 23 (L)	+
2	Superior frontal gyrus 21 (L)	+
3	Superior frontal gyrus 8 (R)	+
4	Superior frontal gyrus 12 (L)	+
5	Frontal pole (L)	-
6	Superior frontal gyrus 4 (L)	+
7	Rostral middle frontal gyrus 10 (L)	+
8	Superior frontal gyrus 15 (R)	+
9	Superior frontal gyrus 9 (L)	+
10	Rostral middle frontal gyrus 1 (L)	-
11	Superior frontal gyrus 39 (L)	+
12	Rostral middle frontal gyrus 25 (L)	-
13	Superior frontal gyrus 37 (R)	+
14	Rostral middle frontal gyrus 1 (R)	-
15	Superior frontal gyrus 19 (L)	+
16	Precentral gyrus (R)	+
17	Postcentral gyrus 15 (R)	+
18	Postcentral gyrus 10 (R)	+
19	Rostral middle frontal gyrus 26 (R)	-
20	Rostral middle frontal gyrus 5 (L)	+

5.4 Connectivity features

The second type of features that will be explored are connectivity features. These features reflect differences in functional connectivity between the healthy controls and the depression group. Functional connectivity is a general term describing different techniques that evaluate the functional connection between different brain regions that might be spatially separated [4], [59]. The most common connectivity measures are correlation, cross-correlation (CR), coherence, granger causality (GCI), directed transfer function (DTF), partial directed coherence (PDC) and mutual information (MI) [4].

5.4.1 Functional connectivity measures

The functional connectivity measures differ from each other and thus reflect different aspects of the functional connection. The properties most commonly used to distinguish the measures are whether they are directed or undirected, whether they are bi- or multivariate, whether they operate in the time- or frequency domain and whether they are linear or non-linear. The properties of the previously mentioned connectivity measures are shown in table 5.3

Undirected vs directed

An undirected connectivity measure can only test whether a functional connection exists between brain regions. Directed connectivity measures can, aside from the validation of a possible connection, also determine in which way the connection exists. This direction in the connection gives extra information as it appoints a role to the different brain regions: one brain region has influenced the other brain region. The direction of a functional connection can be used to explore how signals propagate in the brain.

Bivariate vs multivariate

Bivariate connectivity measures can only analyze the connection of two different brain regions at a time. More complex relationships between multiple brain regions (e.g. brain networks such as the default mode network) can be documented using multivariate connectivity measures.

Time domain vs frequency domain

Time domain connectivity measures will analyze if a connection exists between brain regions in the time domain while frequency domain measures analyze possible connections in the frequency domain.

Linear vs non-linear

Linear connectivity measures will analyze if a linear relationship exists between brain regions while the non-linear connectivity measures also analyze if non-linear relationships are present.

Table 5.3: Functional connectivity measures [4].

Measure	Un-/directed	Bi-/multivariate	Time/Frequency	Non-/Linear
Correlation	undirected	bivariate	time	linear
CR	directed	bivariate	time	linear
Coherence	undirected	bivariate	frequency	linear
GCI	directed	bivariate	time	linear
DTF	directed	multivariate	frequency	linear
PDC	directed	multivariate	frequency	linear
MI	undirected	bivariate	time	non-linear

5.4.1.1 Selection process of connectivity measures

The selection of connectivity measures is, as explained in section 5.4.1, best done using the distinguishing properties. Aside from the properties it is important to keep in mind what the final goal of the connectivity measures is: the classification of people in two possible groups. Using this consideration, a conclusion can be drawn which is that the final result of a connectivity measure should be a single value, or should be easily reducible to a single value, as this is needed for classification.

Frequency based connectivity measures are not investigated. Due to the limited frequency range ($[\frac{1}{128}$, 0.5] Hz) imposed by the high frequency filter (see section 5.1.6) and the limitation of the EPI sequence (see section 2.1.3.1), the frequency spectrum of fMRI is of little interest for feature selection when compared to EEG data.

A second reduction is done by eliminating all directed features. This reduction is defended by the fact that fMRI measures the hemodynamic response, which has variable time delays throughout the brain. Due to this variability, directed connectivity measures are difficult to interpret. A second validation for this reduction is the fact that previous attempts at depression classification based on resting state fMRI used undirected methods [30], [31].

Using both reductions, two connectivity measures are selected: correlation and mutual information. Both will be explained in detail below.

Correlation

Correlation, the Pearson correlation coefficient more precisely, is a measure of the linear correlation between two signals. It has a possible range of [-1,1] where -1 means there is a total negative correlation between both signals, 1 means there is a total positive correlation between both signals and 0 means no correlation exists between the signals [60]. Although correlation might seem a simple connectivity measure, it is proven that it can perform equally as well as mutual information in correctly defining linear connections and is commonly used in functional connectivity studies [30], [31], [61].

The calculation of the Pearson correlation coefficient of two signals, A and B, is shown in formula 5.3. Here N denotes the amount of different values both signals A and B consist of, A_i, μ_A and B_i, μ_B are the ith sample and mean value of variable A and B respectively. The MATLAB function *corrcoef* was used to calculate the Pearson correlation coefficient.[27]

$$\rho(A,B) = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{A_i-\mu_A}{\sigma_A}\right)\left(\frac{B_i-\mu_B}{\sigma_B}\right) \tag{5.3}$$

Mutual information

Mutual information is a non-linear connectivity measure which quantifies the mutual dependency of two signals. It calculates how much information can be obtained of the second variable when only the first variable is observed [62]. The possible range of values is [0,+∞[, where a higher value denotes more shared information between both signals. Because this range is unbounded, the only information one can extract is the difference between two values [62]. Formula 5.4 represents the calculation process. Here $p_{AB}(a,b)$ denotes the joint probability between signal A and B, calculated using their combined histogram; p_A and p_B denote the probability of signal A and B respectively,

27. https://www.mathworks.com/help/matlab/ref/corrcoef.html

these can again be calculated using their histograms. A MATLAB package[28] is used to calculate the mutual information.

$$MI_{AB} = \sum_{a,b} p_{AB}(a,b) log\left(\frac{p_{AB}(a,b)}{p_A(a)p_B(b)}\right) \tag{5.4}$$

5.4.2 Feature selection process

The feature selection process is shown in figure 5.7. It is quite similar to the feature selection process of the intensity-based features (see section 5.3). It starts with the preprocessed data, as described in section 5.1, and can be divided in two parts: a patient specific part and a group specific part. The patient specific part of the feature selection process contains the preparation of the data for each patient while the group specific part contains the steps performed on a group level (depression patients and healthy controls). Each step will be explained.

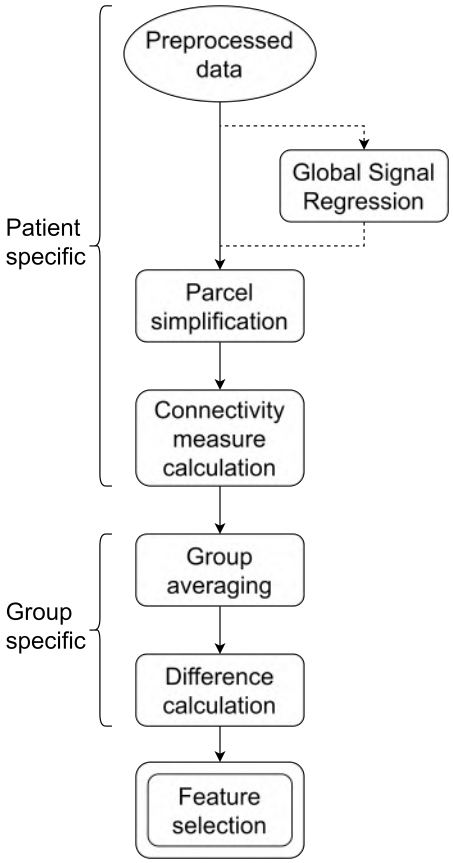


Figure 5.7: The connectivity feature selection process.

5.4.2.1 Global signal regression

The principle of global signal regression is the idea that certain processes, such as cardiac pulse and breathing, influence the BOLD signals (see section 2.2.1) captured by the MRI scanner. These processes will influence the BOLD signal throughout the whole brain by introducing an extra signal that is common in every time series captured. The captured signal is the addition of two signals: a global signal and the local signal caused by local neuronal changes. The addition of a common signal in

28. https://www.mathworks.com/matlabcentral/fileexchange/13289-fast-mutual-information-of-two-images-or-signals

every time series leads to an artificially increased connectivity between every voxel. This artificial increase leads to falsely inflated connectivity values. This increase in connectivity throughout the brain is unwanted as only the neuronal changes within the brain are a true indication of connectivity between brain regions.

To counter the introduction of a global signal, global signal regression is applied. Global signal regression calculates the global signal by averaging all time series and uses the average signal as a regressor in a general linear model. The general linear model regresses the global signal out of the time series, which are now assumed to only contain the information regarding the neuronal changes within the brain. As only the neuronal changes are present in the time series, connectivity values now reflect true connectivity between brain regions [63], [64].

Global signal regression is however subject of much discussion within the field of neuroscience. The principal reasoning behind global signal regression is correct, but an assumption is made. The assumption is the idea that the influences inducing the global signal and the local neuronal changes are completely unrelated (they are orthogonal to each other), and that by regressing the global signal out, one does not regress part of the neuronal signal out. Multiple studies have concluded that global signal regression actually introduces anti-correlated networks and therefore might itself introduce false correlations [65], [66], [67]. This makes the use of global signal regression a controversial choice.

As no definite conclusion is reached regarding global signal regression, connectivity features will be calculated with and without global signal regression as a preprocessing step (as indicated by the dotted line in figure 5.7). The data set created without the global signal regression step will be called the non-regressed data set, the data set created with the global signal regression step will be call the regressed data set. The CONN toolbox is used to perform global signal regression [38].

5.4.2.2 Parcel simplification

A second step that is performed is parcel simplification. This step will reduce all time series from a parcel into a single time series. This step is necessary for two reasons: dimension reduction and intrapatient/intergroup voxel variations.

Dimension reduction is needed as every patient has 155648 (64x64x38) unique time series, 30076 time series which are overlapping with the Lausanne brain atlas [50]. If no dimension reduction is applied, a connectivity measure matrix with size 30076x30076 is obtained, which is computationally not achievable.

The intrapatient/intergroup variations, meaning voxel variability both within a single patient scan and between different patient scans, make single voxel time series highly unreliable as a basis for features.

The technique used to perform the parcel simplification is principal component analysis (PCA). PCA is a statistical technique in which a group of possible correlated observations (here time series) are translated into a set of new variables that are not correlated anymore (they are orthogonal to each other). Each variable in the new set is called a principal component [68]. Only the first principal component of each parcel will be used to calculate functional connectivity between parcels as it will compensate for as much variance as possible.

The result from the parcel simplification step is a 1xN matrix for each patient, where N is 83, 129, 234, 463 or 1015 depending on the chosen atlas level (see section 5.2.1). Each cell contains the first

principal component of the corresponding parcel. Principal component analysis is performed using a MATLAB toolbox.[29]

5.4.2.3 Connectivity measure calculation

The next step in the feature selection process is the connectivity measure calculation. The two connectivity measures that are used are correlation and mutual information (see section 5.4.1.1). The result of the connectivity measure calculation is an NxN matrix for each patient, where N is 83, 129, 234, 463 or 1015 depending on the chosen atlas level (see section 5.2.1). Cell $C_{i,j}$ contains the calculated connectivity between parcels *i* and *j*. These matrices are symmetric because undirected connectivity measures are used, meaning that $C_{i,j}$ is equal to $C_{j,i}$. It should be noted that mutual information for N = 463, 1015 is not calculated as the computation time is too long.

5.4.2.4 Group averaging

The next step is group averaging. An average matrix is calculated for both groups (depression patients and healthy controls). The result of this step is a group matrix for each group, for each atlas level and for each connectivity measure.

5.4.2.5 Difference calculation

After the calculation of the group matrices, the depression group matrices are subtracted from the healthy controls group matrices in order to obtain the difference matrices.

5.4.2.6 Feature selection

The last step is the selection of the features. This is done by finding the 20 biggest values for each difference matrix. These values reflect the connectivity measures that on average differ the most between both groups and thus will be used to differentiate between both groups. The significance of the features is validated statistically using the two-sample t test, done by the *ttest2* function from MATLAB. The features found for mutual information calculated from the non-regressed data set are not significant. All possible features from this feature subtype are tested using the t-test, none are statistically significant. Classifiers trained on this feature set do not perform well (see section 7.3.1).

As five atlas levels for correlation and three atlas levels for mutual information are used and two data sets are used (regressed and non-regressed), sixteen feature sets are obtained.

29. https://www.mathworks.com/matlabcentral/fileexchange/38300-pca-and-ica-package

5.4.3 Interpretation of the found features

A total of 16 feature sets, each containing 20 features, have been calculated. Four sets will be discussed in this section: the best performing feature sets of each connectivity measure for both the regressed and non-regressed data set. The results are shown in section 7.3.1 and discussed in secton **??**.

Interpretation of the used tables

The *number* column defines the significance of the feature. Feature 1 defines the connection between two parcels that has the biggest value in the difference matrix (see section 5.4.2.5), meaning that this connection has on average the highest difference between the healthy controls and the depression group, feature 20 the twentieth biggest difference. The *region1* and *region2* columns define the two parcels that are involved in the connection the feature reflects. "(L)" denotes that the parcel is located in the left hemisphere, "(R)" denotes the right hemisphere. The number at the end of the region name is its further subdivision number (brain regions are subdivided in smaller regions when higher atlas levels are used). The *sign* column defines the sign of the difference value. A "+" denotes a positive difference value, meaning that the connection between the two parcels is higher with the healthy controls when compared to the depression group. A "-" sign denotes the contrary.

5.4.3.1 Features calculated with the non-regressed data set

Correlation

The twenty most significant features are shown in table 5.4. Feature 13, 16 and 20 are closely linked to depression. The connections reflected by feature 13 and 16 both contain the anterior cingulate cortex (ACC), which is an important factor in several models of depression [69], [70], feature 20 contains the orbitofrontal cortex, which has been discussed in section 5.3.3.2. As this feature set did not perform as well as the correlation with regressed data feature set, it will not be discussed further. The results can be found in section 7.3.1, the discussion of the results in section **??**.

Table 5.4: Correlation with non-regressed data set features (Atlas2).

Number	Region1	Region2	Sign
1	Inferior temporal gyrus (L)	Middle temporal gyrus (L)	+
2	Superior frontal gyrus (R)	Precentral gyrus (L)	+
3	Precentral gyrus (L)	Postcentral gyrus (L)	+
4	Caudal middle fontal gyrus (R)	Rostral middle frontal gyrus (R)	+
5	Precentral gyrus (R)	Inferior temporal gyrus (R)	+
6	Precentral gyrus (R)	Paracentral gyrus (L)	+
7	Middle temporal gyrus (R)	Banks of the superior temporal sulcus (L)	-
8	Pars triangularis (L)	Pars opecuaris (L)	+
9	Pericalcarine (R)	Lingual gyrus (L)	-
10	Postcentral gyrus (R)	Superior temporal gyrus (R)	+
11	Lateral orbitofrontal cortex (L)	Insula (Left)	+
12	Precentral gyrus (R)	Superior temporal gyrus (R)	+
13	Lateral orbitofrontal cortex (L)	Caudal anterior cingulate cortex (L)	+
14	Superior temporal gyrus 1 (R)	Superior temporal gyrus 3 (R)	-
15	Caudal middle fontal gyrus (R)	Precentral gyrus (R)	+
16	Rostral anterior cingulate cortex (R)	Lateral orbitofrontal cortex (L)	+
17	Superior frontal gyrus (R)	Pars opercularis (L)	+
18	Precentral gyrus 3 (L)	Precentral gyrus 1 (L)	+
19	Precentral gyrus (R)	Supramarginal gyrus (R)	+
20	Lateral orbitofrontal cortex 3 (R)	Lateral orbitofrontal cortex 1 (R)	-

Mutual information

The twenty most significant features are shown in table 5.5. All features reflect a connection in which the right or left superior frontal gyrus is involved. Seven features (feature 2, 7, 8, 11, 13, 16 an 18) reflect a connection within the left or right superior frontal cortex itself. The accuracy results of this feature set are lower than 50% (see section 7.3.1.7), so these features will not be discussed further.

Table 5.5: Mutual information with non-regressed data set features (Atlas3).

Number	Region 1	Region 2	Sign
1	Superior frontal gyrus 6 (R)	Superior frontal gyrus 8 (L)	-
2	Superior frontal gyrus 6 (R)	Superior frontal gyrus 7 (R)	-
3	Superior frontal gyrus 8 (R)	Superior frontal gyrus 8 (L)	-
4	Superior frontal gyrus 8 (L)	Superior frontal gyrus 7 (R)	-
5	Superior temporal gyrus 4 (R)	Superior frontal gyrus 8 (L)	-
6	Superior frontal gyrus 6 (R)	Lateral orbitofrontal cortex (R)	-
7	Superior frontal gyrus 6 (R)	Superior frontal gyrus 8 (R)	-
8	Superior frontal gyrus 8 (L)	Superior frontal gyrus 6 (L)	-
9	Caudal middle frontal gyrus 3 (L)	Superior frontal gyrus 8 (L)	-
10	Superior frontal gyrus 6 (R)	Caudal middle frontal gyrus 3 (L)	-
11	Superior frontal gyrus 7 (L)	Superior frontal gyrus 8 (L)	-
12	Superior frontal gyrus 6 (R)	Superior frontal gyrus 6 (L)	-
13	Superior frontal gyrus 6 (R)	Superior frontal gyrus 4 (R)	-
14	Superior frontal gyrus 8 (L)	Lateral occipital sulcus 1 (L)	-
15	Superior frontal gyrus 8 (L)	Lateral orbitofrontal cortex (R)	-
16	Superior frontal gyrus 4 (R)	Superior frontal gyrus 8 (L)	-
17	Superior frontal gyrus 6 (R)	Superior temporal gyrus 4 (R)	-
18	Superior frontal gyrus 8 (R)	Superior frontal gyrus 7 (R)	-
19	Superior frontal gyrus 6 (R)	Lateral occipital cortex 1 (L)	-
20	Rostral middle frontal gyrus 5 (L)	Superior frontal gyrus 8 (L)	-

5.4.3.2 Features calculated with the regressed data set

Correlation

The twenty most significant features are shown in table 5.6. Several features are linked to depression, each significant feature will be discussed.

Feature 2 reflects the connection between the right medial orbitofrontal cortex and the left anterior cingulate cortex. Connections between the frontal cortex and the anterior cingulate cortex are disturbed in depression [55], [71]. Feature 5 reflects the connection between the left and right precuneus. The precuneus is a part of the default mode network (DMN) and disturbances in the default mode network have been reported in people with depression [72], disturbances in this specific connection between the left and right precuneus in people with depression are however not found in literature. Feature 9 reflects the connection between the precuneus and the isthmus cingulate cortex, which could also be related to depression. No literature for this specific connection has been found. Feature 13 and 15 are two connections in the prefrontal cortex, which is involved in depression [54], [56]. The sign of both connections is negative, which is counter-intuitive as the prefrontal cortex is hypoactive in depression [55]. A possible explanation for this could be that hypoactivity does not mean a decrease in correlation between brain regions. No literature for these specific connections has been found.

Table 5.6: Correlation with regressed data set features (Atlas3).

Number	Region 1	Region 2	Sign
1	Rostral middle frontal gyrus (L)	Supramarginal gyrus (L)	+
2	Medial orbitofrontal cortex (R)	Rostral anterior cingulate cortex (L)	+
3	Fusiform gyrus (R)	Inferior temporal gyrus (R)	-
4	Lingual gyrus (R)	Pericalcarine gyrus (L)	+
5	Precuneus (R)	Precuneus (L)	+
6	Precentral gyrus (R)	Postcentral gyrus (L)	+
7	Paracentral gyrus 1 (L)	Paracentral gyrus 2 (L)	-
8	Precentral gyrus (R)	Precentral gyrus (L)	-
9	Isthmus cingulate cortex (L)	Precuneus (L)	+
10	Fusiform gyrus	Lateral occipital cortex (L)	-
11	Lateral occipital sulcus (L)	Fusiform gyrus (L)	+
12	Pericalcarine gyrus (R)	Pericalcarine gyrus (L)	+
13	Superior frontal gyrus 6 (R)	Superior frontal gyrus 7 (R)	-
14	Superior parietal gyrus (R)	Lateral occipital cortex (R)	+
15	Superior frontal gyrus (R)	Caudal middle frontal gyrus (R)	-
16	Precentral gyrus (R)	Supramarginal gyrus (R)	+
17	Lateral occipital cortex (R)	Fusiform gyrus (R)	+
18	Superior temporal gyrus (R)	Precentral gyrus (L)	-
19	Superior parietal gyrus 3 (L)	Superior parietal gyrus 4 (L)	-
20	Cuneus (R)	Cuneus (L)	-

Mutual information

The twenty most significant features are shown in 5.7. A first observation is that a few brain regions return in several features. The four brain regions present in multiple features are the right paracentral lobule, the right inferior parietal gyrus, the brainstem and the right superior parietal gyrus. The right paracentral lobule (RPL) is present in 11 features (feature 1, 2, 3, 7, 8, 9, 10, 13, 16, 18 and 19), the right inferior parietal gyrus (RIPG) is present in 9 features (feature 2, 4, 5, 6, 7, 12, 15, 17 and 20), the right superior parietal gyrus (RSPG) is present in four features (feature 1, 5, 11 and 14) and the brainstem is present in three features (feature 3, 4 and 11).

Changes in connectivity in the paracentral lobule in people with depression have been described in literature [73], [74], [75]. No clear explanation of its involvement in depression is given and different literature also reports both increased and decreased functional connectivity within the region. Further research needs to be done to investigate this region. The right inferior and superior parietal lobe have also been linked with depression, however decreased functional connectivity is reported while increased connectivity is found here [76]. The brainstem contains multiple nuclei that could be involved in depression and research tries to understand its involvement in the disease. The specific connections involving the brainstem however are not found in literature.

The assessment of the clinical relevance of connectivity features is more difficult than the assessment of the intensity or structural features as two regions are involved, making a possible recurrence of the specific connection in literature less likely. An attempt to counter this problem is the dismissal of the subdivision number when possible (subdivision numbers are still shown when a connection within a single region is found), but this did not improve the search for literature significantly. As different connectivity measures are used, a recurrence of a connection in literature is not always proof of its clinical significance. Some skepticism toward the found literature is also necessary; this skepticism is needed for every found relevance for every feature type and is not bounded to the connectivity features.

Table 5.7: Mutual information with regressed data set features (Atlas3).

Number	Region 1	Region 2	Sign
1	Paracentral lobule (R)	Superior parietal gyrus (R)	-
2	Inferior parietal gyrus (R)	Paracentral lobule (R)	-
3	Paracentral lobule (R)	Brainstem	-
4	Brainstem	Inferior parietal gyrus (R)	-
5	Inferior parietal gyrus 5 (R)	Superior parietal gyrus (R)	-
6	Inferior parietal gyrus 1 (R)	Inferior parietal gyrus 5 (R)	-
7	Inferior parietal gyrus 1 (R)	Paracentral lobule (R)	-
8	Paracentral lobule (R)	Parahippocampal gyrus (L)	-
9	Paracentral lobule (R)	Supramarginal gyrus (R)	-
10	Rostral middle frontal gyrus (R)	Paracentral lobule (R)	-
11	Superior parietal gyrus (R)	Brainstem	-
12	Inferior parietal gyrus (R)	Parahippocampal gyrus (L)	-
13	Fusiform gyrus (R)	Paracentral lobule (R)	-
14	Parahippocampal gyrus (L)	Superior parietal gyrus (R)	-
15	Inferior parietal gyrus (R)	Lateral occipital cortex (L)	-
16	Paracentral lobule (R)	Superior temporal gyrus (L)	-
17	Supramarginal gyrus (R)	Inferior parietal gyrus (R)	-
18	Rostral middle frontal gyrus (L)	Paracentral lobule (R)	-
19	Lateral occipital cortex (L)	Paracentral lobule (R)	-
20	Inferior parietal gyrus (R)	Fusiform gyrus (R)	-

Chapter 6

Classifier training

A pipeline is designed to train each classifier, this to be able to compare results between different feature sets (see chapter **??**). The complete classification training pipeline is shown in figure 6.1. Every step of this pipeline will be explained.

6.1 Classification training pipeline

6.1.1 Starting point

The starting point of the classifier training pipeline is a feature set, the final result from one of the feature selection processes (see chapter **??**). The feature set is a 106x(N+1) matrix, where N equals the amount of features. The number 106 denotes the amount of people that are present in the whole data set (60 healthy controls and 46 patients with depression). The last column of the feature set contains the class of each patient. One denotes the *healthy control* class, minus one the *depression* class.

6.1.2 Class balancing

The size of both classes is unequal. This can pose a problem in the training of a classifier because a bias could be introduced towards the more abundant class. If class A is ten times more abundant than class B, a machine learning technique that always predicts class A, even without being trained, will have an accuracy of over 90%; this is called the accuracy paradox. Two main possibilities exist to counter this problem: class imbalance learning methods and class balancing. Class imbalance learning methods refer to adaptions of learning methods that can counter the problem of class imbalance [88], [89].

Class balancing refers to the dismissal of instances from one class so that both classes are represented equally. This method has a disadvantage when compared to class imbalance learning methods: the dismissal of instances makes the training set smaller, which might lead to overfitting[30]. Furthermore is class balancing not possible if the imbalance between both classes is too big.

Class balancing is chosen as a balancing method because the dismissed persons (fourteen healthy controls) can be used as a second validation set (see section 6.1.4.2). The second validation set can be used to check if the trained model is not overfit. Each time the classification training pipeline is used, the fourteen healthy controls that will not be used for the training of the classifier are chosen by random permutation. This ensures that all healthy controls can be used in the training of the classifier.

30. https://www.investopedia.com/terms/o/overfitting.asp. Overfitting: a modeling error which occurs when a function is too closely fit to a limited set of data points.

6.1.3 Train-validation splitting

The 46 healthy controls that are selected are, together with the 46 patients with depression, the data set (92 people in total) that is used to train and validate the classifier. The first step is to split this data set in a training set and a validation set. A common choice is the 80-20 split [44]. 80 percent of the total data set (73 people) are selected by a random permutation to be the training set, 20 percent (19 people) are selected to be the validation set.

6.1.4 Model training and validation

The final step in the classification training pipeline is model training and model validation.

6.1.4.1 Training

The support vector machine (see chapter 3) is trained using the training set.

6.1.4.2 Validation

The final step in the classification training pipeline is validation. The trained model will be validated using the validation set defined in section 6.1.3. The features of the people in the validation set will be given to the trained classifier as input. Afterwards the classifier will predict whether the given input features belong to a person from the healthy control group or from the patients with depression group. The validation is performed to know if the trained classifier is adequate. A second validation set is used, it contains the fourteen people dismissed in section 6.1.2.

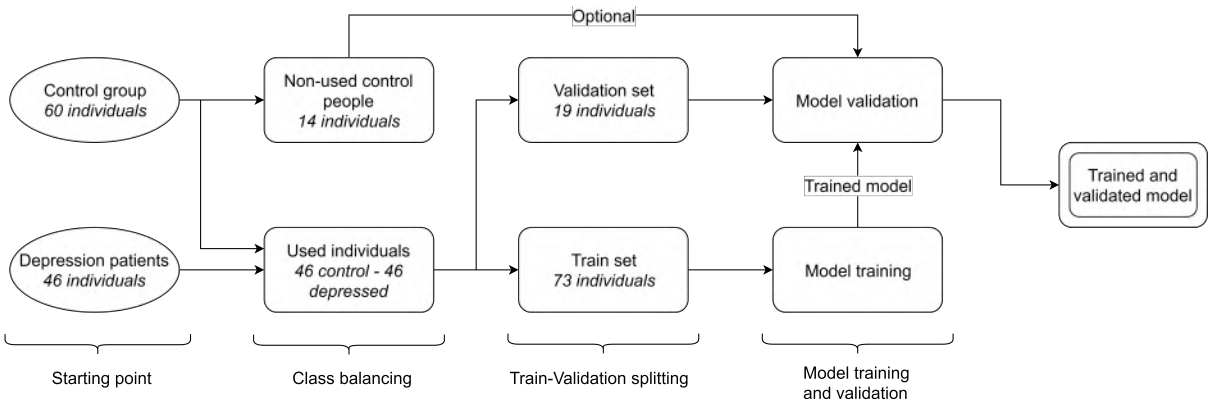


Figure 6.1: The complete classification training pipeline.

Chapter 7

Results

In this chapter the found results of the different trained classifiers will be discussed. The first part of the chapter will present the results of the best performing classifiers (this is called the global results). The second part presents the results of each feature subtype classifier seperately. The third part presents the results of the combined feature classifiers Firstly the used method to retrieve data as well as the visual representation of the results are explained.

7.1 Data collection and presentation

7.1.1 Data collection

Three different feature selection methods are used: the first method lets the algorithm itself choose how many features are optimal for classification, the second method again lets the algorithm itself choose which features are best for classification, but imposes the demand that 10 features need to be chosen, the third method uses all 20 features of a feature set. Twenty iterations of each feature selection method are used in the machine learning pipeline (see section 6.1) to obtain an average accuracy of each method. This reduces the chances of falsely inflated or deflated results for the feature set. The final trained classifier is validated twice: firstly with the validation set defined by the train-validation splitting step (see section 6.1.3) and secondly with the non-used control group defined in the class balancing step (see section 6.1.2). The second validation set will be called the *optional validation set* from now on.

The classification pipeline (see section 6.1) will be used 60 times for each feature set. Exceptions to this are the right hemisphere feature set as it only consists of 11 features (this feature set will be used 40 times: once with the variable amount of features, once with all 11 features) and the parcel volume feature set as it only consists of 6 features (this feature set will be used 20 times with all 6 features). The combined feature sets, shown in section 7.3.2, are used 20 times with all features. Both the intensity and connectivity feature sets contain 40 features, the structural feature set contains 36 features (19 left hemisphere features, 11 right hemisphere features and 6 parcel volume features).

7.1.2 Presentation

Each classification result will be presented by a table showing the accuracy of both the validation set (19 people, "Val." in the table) as well as the optional validation set (14 people, "Opt." in the table), the numbers are percentages. The mean accuracy and standard deviation of each result distribution are shown at the bottom of the table. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the validation set are also shown. These are not calculated for the

optional data set because it is impossible to have a true positive and false negative value. A graph presents the obtained results as a violin plot (see section 7.1.2.1).

7.1.2.1 Violin plot

A violin plot is an extension of a box plot. It shows beside the mean, median and outliers also the probability density function of the distribution it represents. The thickness of a violin plot at point x represents the probability of x given the represented distribution.[31] The violin plots are created using a MATLAB toolbox.[32] It should be noted that a simplified version of the violin plot is used; it only shows the mean and the probability density function. Each figure will contain at least two violin plots: one blue and one yellow. Blue violin plots show the results of the validation set (19 people), yellow violin plots show the results of the optional validation set (14 people) (see chapter 6).

Violin plots are used as they clearly show differences between both classification distributions as well as the compactness/variability of each distribution. The compactness of a distribution is also represented by the standard deviation in the table. Differences in classification distributions between the validation and optional validation set could be a sign of high variability within the feature set or of overfitting. A slightly lower compactness of the optional validation set is however expected compared to the validation set, as it contains less people. A low compactness of a result distribution is a sign of a feature set that is not optimal for classification, as the choice of test set and start conditions of the training is the defining factor contrary to the quality of the features itself. High accuracy classifiers in a low compactness result distribution should only be considered as viable for classification if the accuracy of the iteration is high in both the validation and the optional validation set.

An example of possible feature results is shown in figure 7.1. The two violin plots on the left show the result distributions of a high quality feature set. Both have high mean accuracies (denoted by the black line) and have high compactness as they are small violin plots. No significant difference in mean accuracy between both result distributions again shows the consistent ability of the classifiers to correctly diagnose depression. The two center violin plots show a bad result as the left violin plot has a high mean accuracy and a high compactness while the right violin plot has a low mean accuracy and a low compactness. The big difference between both violin plots show that the classifiers are not able to consistently predict depression. The two violin plots on the right show again a bad result as both result distributions have low mean accuracies and low compactness.

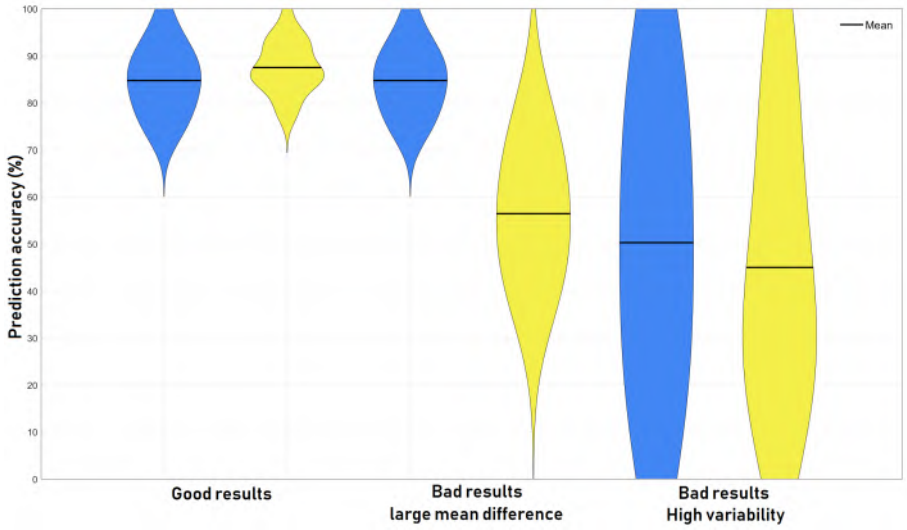


Figure 7.1: Example of the different possible results and violin plots.

31. https://datavizcatalogue.com/methods/violin_plot.html
32. https://www.mathworks.com/matlabcentral/fileexchange/45134-violin-plot

7.2 Global results

The global results represent the best performing classifier of each feature subtype. The results are shown in figure 7.2. Six feature subtype result distributions are shown: the left hemisphere thickness, parcel volume, absolute intensity (atlas3), relative intensity (atlas3), mutual information with regressed data (atlas3) and correlation with regressed data (atlas3) result distribution. The result distributions that are shown are from the validation set and not the optional validation set (hence the color of every result distribution). These specific result distributions are chosen as they are the best single feature subtype result distributions (with respect to the mean accuracy and standard deviation).

Two conclusions from the global results can be drawn. The first is that the mean accuracy of the structural features (left hemisphere thickness and parcel volume) is much lower than the mean accuracy of both the intensity and connectivity features. This could possibly be explained by the fact that the assumption is made that all people in the depression group have had a similar form of depression while in reality this is not true (this is discussed more in depth in section 8.2). A second conclusion is that while the mean accuracies of the intensity (absolute and relative intensity) and connectivity (mutual information with regressed data and correlation with regressed data) result distributions are close together, the connectivity result distributions are slightly higher. The higher compactness of these result distributions also indicate a better feature quality. This could be explained by the fact that the connectivity features extract more information from the fMRI data (this is discussed more in depth in section 8.2).

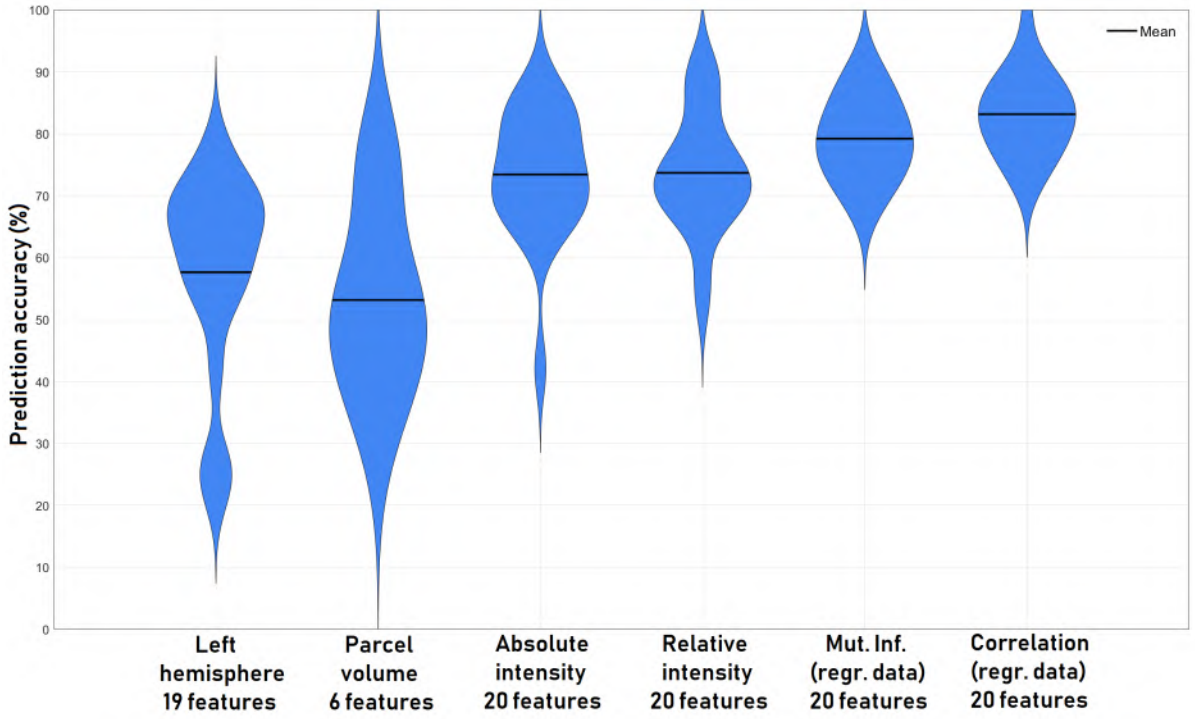


Figure 7.2: Violin plot of the global results.

7.3 Results

The results are discussed for each feature subtype (absolute intensity, relative intensity, correlation with non-regressed data, mutual information with non-regressed data, correlation with regressed data, mutual information with regressed data, left hemisphere thickness, right hemisphere thickness, parcel volume) separately. The results are investigated with respect to the atlas level (for intensity and connectivity feature sets), the amount of features used (for all feature types) and the compactness of the result distributions (for all feature types). Positive outliers are discussed.

7.3.1 Single feature type classifier

7.3.1.1 The left hemisphere thickness classifier

Description

The results of the left hemisphere thickness classifier are shown in table 7.1 and figure 7.3. The mean accuracy remains constant when more features are used for classification, but never reaches levels that could be considered adequate for classification as the best mean accuracy is ±59.5%. Compactness is low for all result distributions. One positive outlier is present: fourth iteration when 10 features are used (average accuracy = ±72%, sensitivity = 0.78, specificity = 0.7, ppv = 0.7, npv = 0.78). It should be noted that this is a positive outlier when only the current result distribution is considered.

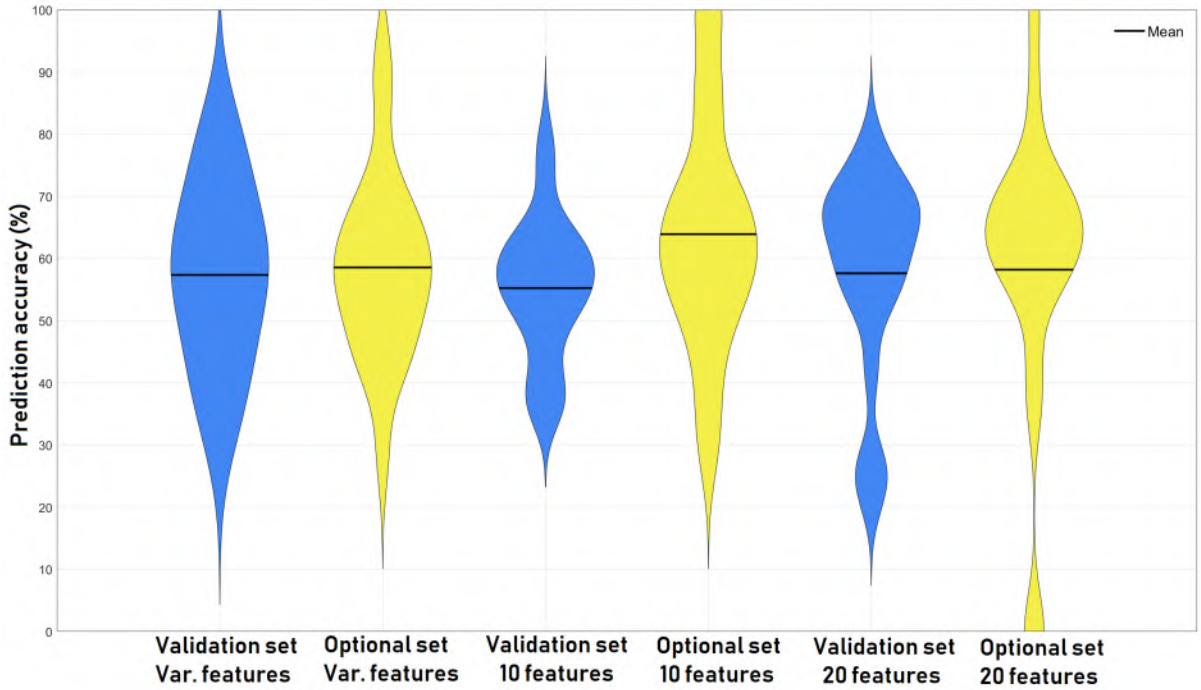


Figure 7.3: Violin plot of the results of the left hemisphere thickness classifier.

Table 7.1: Best results of the left hemisphere thickness classifier.

Iteration	Var. features		10 features		19 features	
	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)
1	63,16	42,86	42,11	100	63,16	64,29
2	78,95	57,14	52,63	64,29	26,32	0
3	68,42	64,29	63,16	71,43	73,68	42,86

4	36,84	28,57	73,68	71,43	21,05	100
5	63,16	71,43	63,16	28,57	26,32	0
6	78,95	57,14	78,95	50	63,16	64,29
7	31,58	57,14	52,63	57,14	78,95	64,29
8	57,89	57,14	52,63	35,71	42,11	71,43
9	57,89	50	47,37	64,29	52,63	85,71
10	78,95	57,14	63,16	64,29	68,42	35,71
11	57,89	42,86	36,84	57,14	52,63	71,43
12	47,37	64,29	57,89	85,71	57,89	71,43
13	42,11	85,71	36,84	85,71	68,42	64,29
14	57,89	64,29	57,89	42,86	57,89	64,29
15	47,37	50	57,89	50	57,89	57,14
16	47,37	92,86	63,16	64,29	68,42	64,29
17	63,16	64,29	36,84	100	68,42	57,14
18	73,68	50	52,63	57,14	68,42	71,43
19	36,84	42,86	57,89	57,14	73,68	57,14
20	57,89	71,43	57,89	71,43	63,16	57,14
Mean accuracy	57,37	58,57	55,26	63,93	57,63	58,21
Standard deviation	14,38	14,95	11,39	18,89	16,6	24
Sensitivity	0.598	-	0.58	-	0.602	-
Specificity	0.549	-	0.525	-	0.551	-
PPV	0.595	-	0.55	-	0.575	-
NPV	0.605	-	0.556	-	0.578	-

7.3.1.2 The right hemisphere thickness classifier

Description

The results of the right hemisphere thickness classifier are shown in table 7.2 and figure 7.4. Only 11 features were statistically significant (see section 4.2.2.2), so the only two different amounts of features were used for classification. The mean accuracy does not increase significantly with respect to the amount of features that are used, the best mean accuracy is reached when a variable amount of features is used (mean accuracy = $\pm$56.5%). Compactness differs significantly between the validation and optional validation set. The reason for this is explained in section **??**. One positive outlier is present: the tenth iteration when a variable amount of features is used (average accuracy = $\pm$82.3%, sensitivity = 0.8, specificity = 0.78, ppv = 0.8, npv = 0.78). It is difficult to know whether this outlier is the product of a good training process or a good starting position.

Table 7.2: Best results of the right hemisphere thickness classifier.

Iteration	Var. features		11 features	
	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)
1	47,37	71,43	36,84	100
2	52,63	57,14	57,89	64,29
3	52,63	35,71	42,11	0
4	52,63	28,57	57,89	42,86
5	52,63	42,86	57,89	50
6	57,89	35,71	57,89	35,71
7	63,16	78,57	73,68	57,14
8	73,68	42,86	57,89	57,14
9	47,37	64,29	63,16	64,29
10	78,95	85,71	47,37	100

11	73,68	42,86	63,16	64,29
12	68,42	71,43	63,16	35,71
13	57,89	28,57	47,37	50
14	47,37	57,14	52,63	57,14
15	57,89	42,86	57,89	50
16	68,42	42,86	57,89	85,71
17	63,16	35,71	36,84	100
18	42,11	57,14	47,37	21,43
19	78,95	64,29	36,84	57,14
20	63,16	64,29	36,84	78,57
Mean accuracy	60	52,5	52,63	58,57
Standard deviation	10,99	16,91	10,66	25,87
Sensitivity	0.626	-	0.554	-
Specificity	0.574	-	0.498	-
PPV	0.595	-	0.535	-
NPV	0.605	-	0.517	-

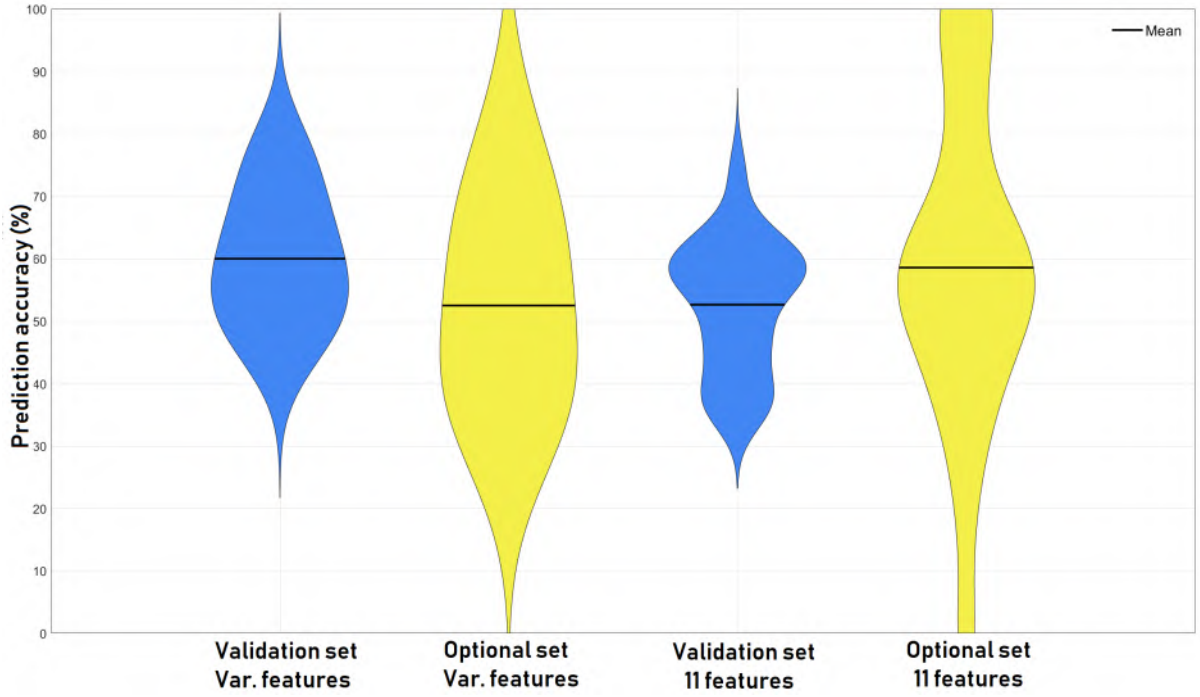


Figure 7.4: Violin plot of the results of the right hemisphere thickness classifier.

7.3.1.3 The parcel volume classifier

Description

The results of the parcel volume classifier are shown in table **??** and figure 7.5. Only 6 features were statistically significant (see section 4.3) so the classifiers will be trained only with all 6 features. The mean accuracy as well as the compactness differ significantly between the validation and optional validation set (mean accuracy = ±61%). One positive outlier is present: the twelfth iteration (average accuracy = ±71.6%, sensitivity = 0.8, specificity = 0.78, ppv = 0.8, npv = 0.78). It should be noted that this is a positive outlier when only the current distribution is considered.

Table 7.3: Best results of the parcel volume classifier

Iteration	6 features	
	Val.(%)	Opt.(%)
1	68,42105	35,71429
2	52,63158	50
3	52,63158	50
4	36,84211	42,85714
5	42,10526	100
6	57,89474	50
7	57,89474	92,85714
8	47,36842	42,85714
9	42,10526	100
10	36,84211	100
11	36,84211	100
12	78,94737	64,28571
13	52,63158	42,85714
14	52,63158	64,28571
15	73,68421	57,14286
16	26,31579	100
17	47,36842	85,71429
18	73,68421	57,14286
19	47,36842	100
20	78,94737	42,85714
Mean accuracy	53,15789	68,92857
Standard deviation	15,07139	24,98926
Sensitivity	0.558	-
Specificity	0.504	-
PPV	0.535	-
NPV	0.528	-

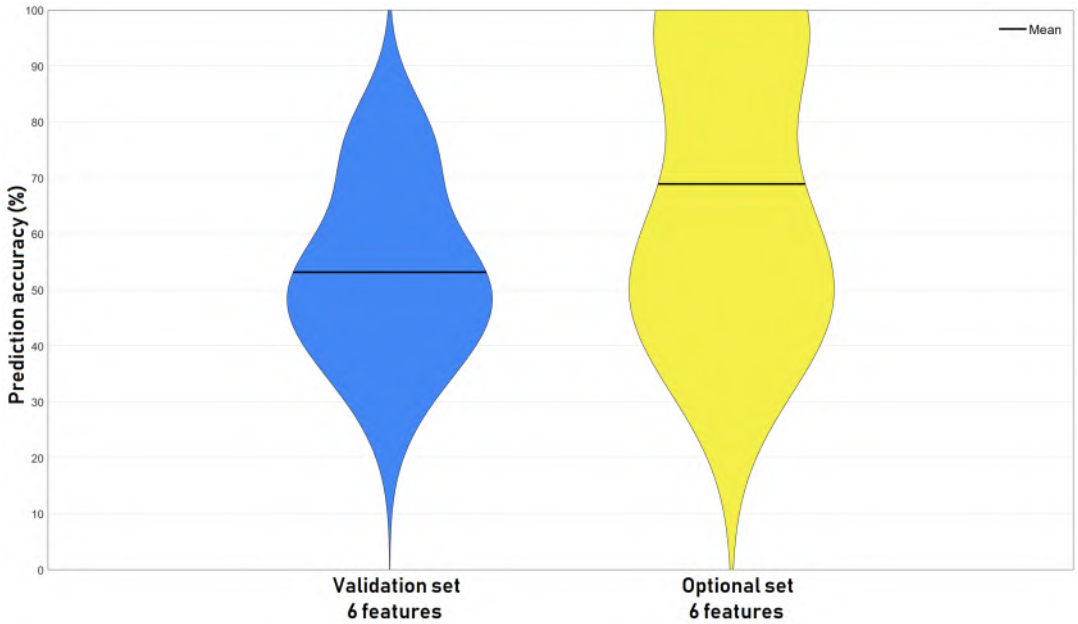


Figure 7.5: Violin plot of the results of the parcel volume classifier.

7.3.1.4 The absolute intensity classifier

Description

The results of the absolute intensity classifier are shown in table 7.4 and figure 7.6. The best average classification accuracy is obtained when the features were calculated using the third atlas level (234 parcels). An increase in classification accuracy is obtained when all 20 features are used for classification compared to the variable amount of features or 10 features. The compactness of the validation set result distributions is similar (SD = ±10.2) for all three feature amounts and is smaller than the compactness of the optional validation sets. A significant increase in compactness and mean accuracy is noticeable for the third optional validation set. Two positive outliers are present: the first iteration when a variable amount of features is used (average accuracy = ±87%, sensitivity = 0.89, specificity = 0.9, PPV = 0.89, NPV = 0.9) and the eleventh iteration when all 20 features are used (average accuracy = ±84%, sensitivity = 0.89, specificity = 0.8, PPV = 0.8, NPV = 0.89).

Table 7.4: Best results of the absolute intensity feature classifier (Atlas3).

Iteration	Var. features		10 features		20 features	
	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)
1	89,47	85,71	73,68	78,57	78,95	78,57
2	63,16	71,43	57,89	64,29	78,95	78,57
3	57,89	57,14	68,42	50	84,21	57,14
4	68,42	42,86	68,42	50	68,42	71,43
5	68,42	78,57	68,42	42,86	42,11	100
6	47,37	71,43	52,63	85,71	73,68	71,43
7	63,16	78,57	68,42	78,57	84,21	78,57
8	63,16	71,43	78,95	64,29	63,16	78,57
9	78,95	35,71	94,74	57,14	84,21	64,29
10	57,89	50	73,68	42,86	73,68	78,57
11	73,68	71,43	57,89	64,29	89,47	78,57
12	68,42	71,43	73,68	78,57	84,21	64,29
13	63,16	50	73,68	78,57	63,16	85,71
14	57,89	64,29	73,68	42,86	68,42	64,29
15	73,68	57,14	47,37	85,71	73,68	78,57
16	47,37	71,43	68,42	85,71	68,42	92,86
17	63,16	42,86	78,95	71,43	73,68	78,57
18	63,16	57,14	68,42	85,71	78,95	78,57
19	47,37	64,29	57,89	71,43	68,42	57,14
20	63,16	14,29	68,42	50	68,42	85,71
Mean accuracy	63,95	60,36	68,68	66,43	73,42	76,07
Standard deviation	10,16	16,82	10,19	15,34	10,32	10,66
Sensitivity	0.67	-	0.712	-	0.762	-
Specificity	0.613	-	0.661	-	0.718	-
PPV	0.635	-	0.685	-	0.72	-
NPV	0.644	-	0.689	-	0.75	-

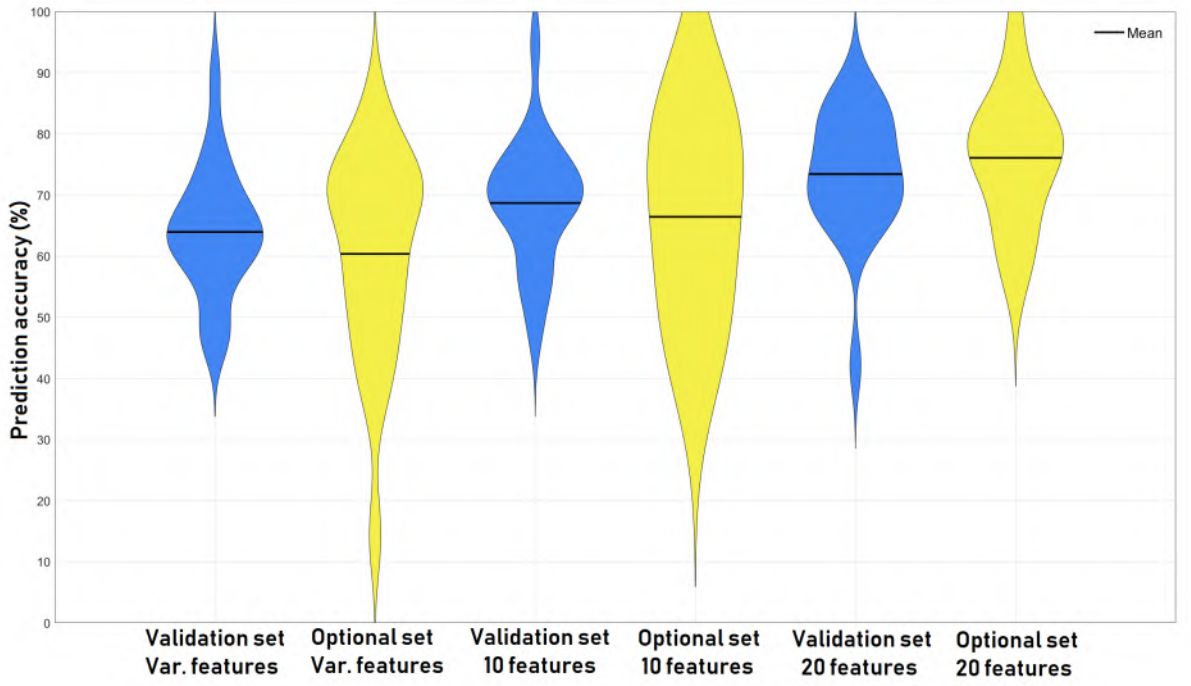


Figure 7.6: Violin plot of the results of the absolute intensity classifier.

7.3.1.5 The relative intensity classifier

Desription

The results of the relative intensity classifier are shown in table 7.5 and figure 7.7. The best average classification accuracy is obtained when the features were calculated using the fifth atlas level (1015 parcels). An increase in classification accuracy is obtained when all 20 features are used for classification compared to the variable amount of features or 10 features (mean accuracy = ±73%). The compactness of the validation set result distributions varies slightly and is lowest when 10 features are used. No significant increase in compactness is noticed when all 20 features are used. Three positive outliers are present: the sixth iteration when a variable amount of features is used (mean accuracy = ±81.5%, sensitivity = 0.778, specificity = 0.8, PPV = 0.8, NPV = 0.778), the thirteenth iteration when all features are used (average accuracy = ±84%, sensitivity = 0.9, specificity = 0.889, PPV = 0.889, NPV = 0.9) and the sixteenth iteration when 10 features are used (average accuracy = ±88%, sensitivity = 0.889, specificity = 0.8, PPV = 0.889, NPV = 0.8).

Table 7.5: Best results of the relative intensity feature classifier (Atlas5).

Iteration	Var. features		10 features		20 features	
	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)
1	52,63	57,14	63,16	78,57	68,42	64,29
2	57,89	64,29	68,42	100	84,21	71,43
3	57,89	71,43	52,63	50	73,68	42,86
4	57,89	57,14	68,42	42,86	68,42	92,86
5	73,68	64,29	52,63	92,86	73,68	85,71
6	78,95	85,71	73,68	71,43	89,47	71,43
7	63,16	42,86	68,42	64,29	73,68	71,43
8	63,16	71,43	57,89	50	68,42	85,71
9	63,16	35,71	36,84	85,71	78,95	50
10	63,16	64,29	73,68	71,43	57,89	71,43

11	63,16	57,14	73,68	85,71	52,63	64,29
12	63,16	50	57,89	85,71	89,47	57,14
13	57,89	57,14	78,95	64,29	89,47	78,57
14	47,37	50	52,63	71,43	68,42	78,57
15	52,63	57,14	57,89	28,57	73,68	57,14
16	57,89	35,71	84,21	92,86	73,68	85,71
17	73,68	85,71	68,42	64,29	73,68	85,71
18	63,16	92,86	78,95	71,43	68,42	78,57
19	73,68	35,71	57,89	85,71	68,42	78,57
20	52,63	35,71	63,16	64,29	78,95	85,71
Mean accuracy	61,84	58,57	64,47	71,07	73,68	72,86
Standard deviation	7,96	16,69	11,15	17,85	9,415	13,09
Sensitivity	0.648	-	0.671	-	0.761	-
Specificity	0.59	-	0.619	-	0.714	-
PPV	0.605	-	0.64	-	0.73	-
NPV	0.633	-	0.65	-	0.744	-

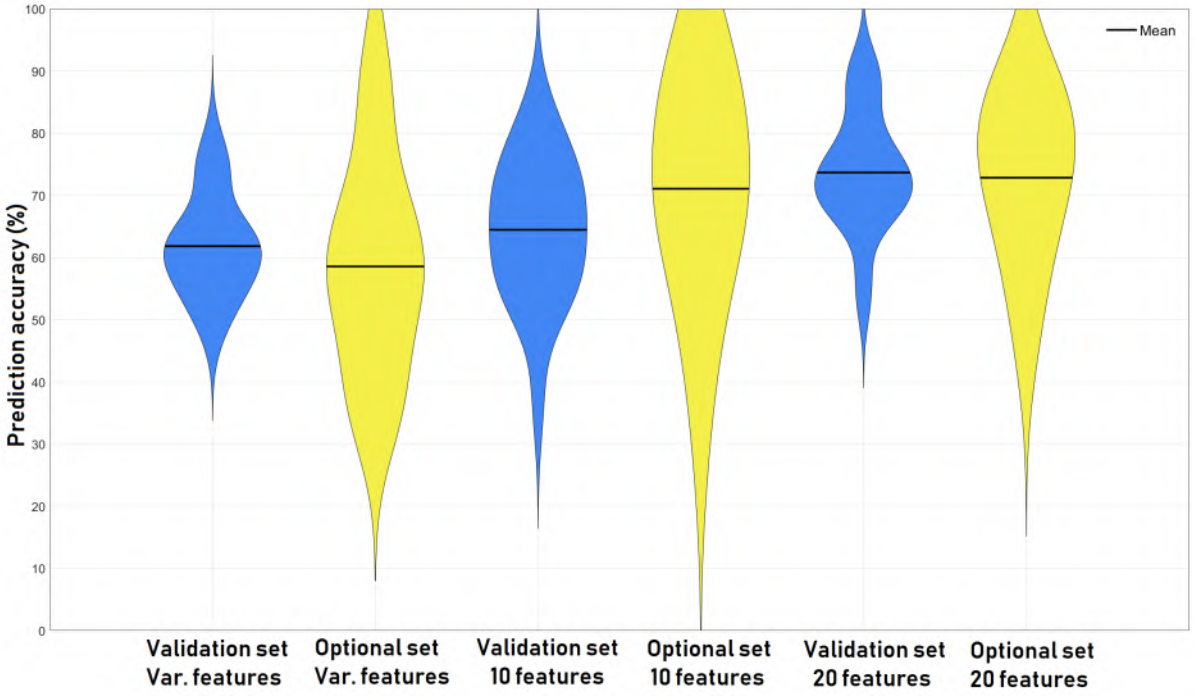


Figure 7.7: Violin plot of the results of the relative intensity classifier.

7.3.1.6 The correlation with non-regressed data classifier

Description

The results of the correlation with non-regressed data classifier are shown in table 7.6 and figure 7.8. The best average classification accuracy is obtained when the features were calculated using the second atlas level (129 parcels). The mean accuracy increases consistently when more features are used, the highest mean accuracy is obtained when all 20 features are used (mean accuracy = ±77%). The compactness of the validation set result distributions are similar (SD = ±10.8), the compactness of the optional validation set fluctuates. The considerable decrease in compactness of the optional validation set result distribution when all 20 features are used (SD = 20.53) is due to both very low

and very high accuracies (fifth and nineteenth iteration, seventh and seventeenth iteration), showing the unreliability of the complete feature set. Two positive outliers are present: the seventh iteration when all 20 features are used (average accuracy = ±86%, sensitivity = 0.8, specificity = 0.778, PPV = 0.8? NPV = 0.778) and the seventeenth iteration when all 20 features are used (average accuracy = ±90.7%, sensitivity = 0.889, specificity = 0.9, PPV = 0.889, NPV = 0.9).

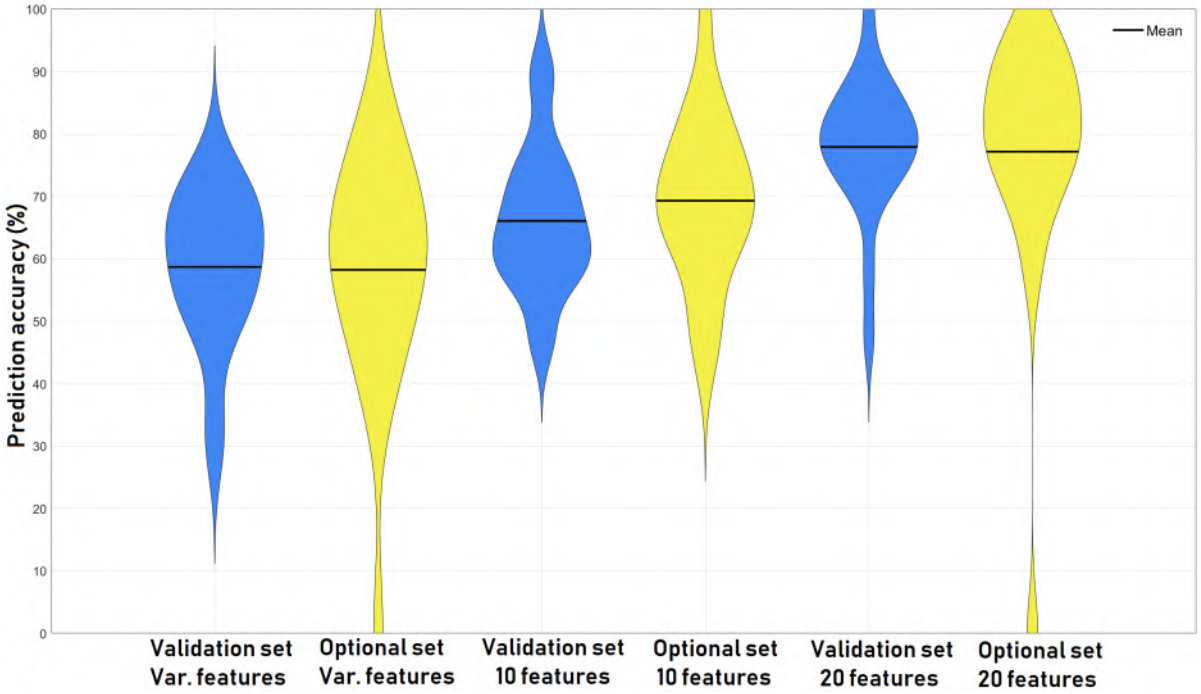


Figure 7.8: Violin plot of the results of the correlation with non-regressed data classifier.

Table 7.6: Best results of the correlation with non-regressed data classifier (Atlas2).

Iteration	Var. features		10 features		20 features	
	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)
1	52,63	64,29	57,89	71,43	68,42	85,71
2	68,42	42,86	63,16	50	84,21	92,86
3	31,58	64,29	63,16	42,86	78,95	78,57
4	47,37	64,29	57,89	85,71	78,95	71,43
5	52,63	71,43	57,89	78,57	47,37	0
6	47,37	0	68,42	78,57	78,95	85,71
7	73,68	42,86	47,37	100	78,95	92,86
8	68,42	57,14	63,16	78,57	84,21	71,43
9	68,42	71,43	78,95	71,43	57,89	100
10	52,63	64,29	73,68	85,71	100	78,57
11	57,89	78,57	89,47	64,29	73,68	92,86
12	52,63	35,71	68,42	71,43	84,21	85,71
13	68,42	50	57,89	64,29	73,68	71,43
14	68,42	78,57	57,89	71,43	73,68	64,29
15	57,89	57,14	73,68	64,29	78,95	78,57
16	63,16	71,43	47,37	57,14	84,21	85,71
17	63,16	85,71	89,47	64,29	89,47	92,86
18	73,68	57,14	73,68	64,29	73,68	78,57
19	63,16	57,14	68,42	71,43	89,47	57,14

20	57,89	50	63,16	50	78,95	78,57
Mean accuracy	59,47	58,21	66,05	69,29	77,89	77,14
Standard deviation	10,27	18,39	11,22	13,19	10,99	20,53
Sensitivity	0.621	-	0.636	-	0.803	-
Specificity	0.569	-	0.685	-	0.756	-
PPV	0.59	-	0.661	-	0.77	-
NPV	0.6	-	0.66	-	0.789	-

7.3.1.7 The mutual information with non-regressed data classifier

Description

The results of the mutual information with non-regressed data classifier are shown in table 7.7 and figure 7.9. The "best" average classification accuracy is obtained when the features were calculated using the third atlas level (234 parcels). The mean accuracy stays stable when more features are used, the highest accuracy is obtained when 10 features are used (mean accuracy = ±49.5%). It should be noted that, as this is a binary classification problem, a random guess between depression or healthy control would result in a higher classification accuracy. The compactness of the validation set result distributions remains consistent (SD = ±10), the compactness of the optional validation set result distributions is very high, especially when all features are used. Only healthy controls are present in the optional validation set (see section 6.1.2), the disjunction in classification accuracies (both very low and very high with no average results) shows the fact that the results are defined by the start position instead of the feature values. No positive outliers are present.

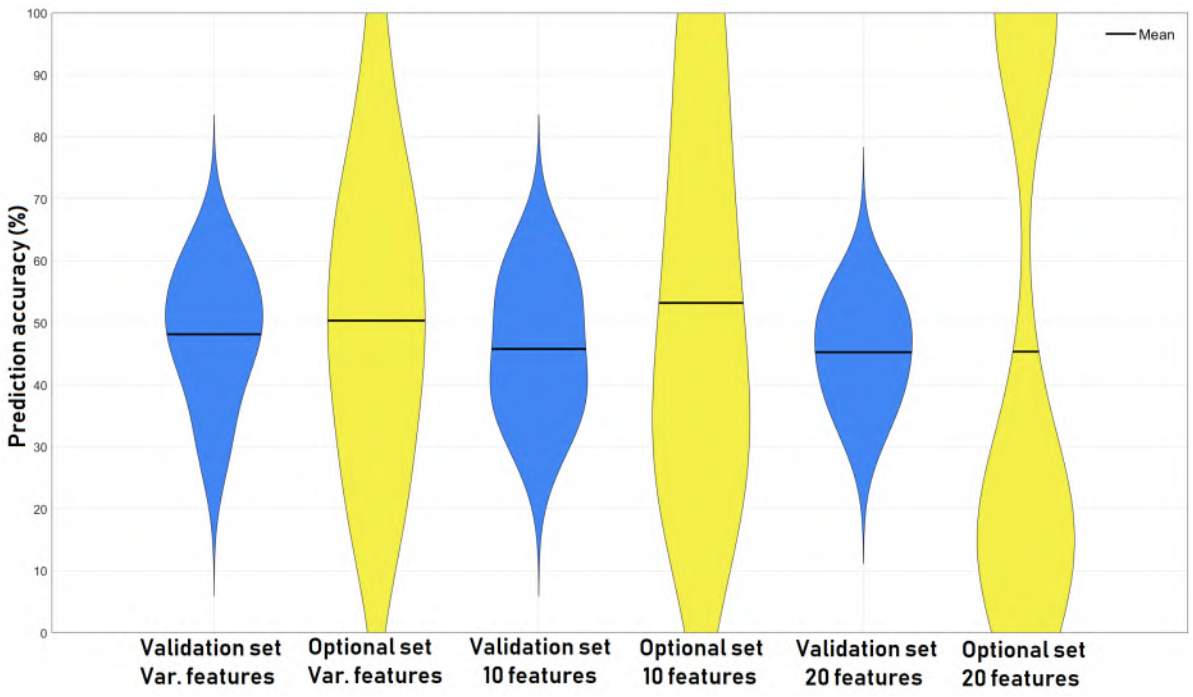


Figure 7.9: Violin plot of the results of the mutual information with non-regressed data classifier.

Table 7.7: Best results of the mutual information with non-regressed data classifier (Atlas3).

Iteration	Var. features		10 features		20 features	
	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)
1	47,37	50	52,63	42,86	52,63	21,43
2	57,89	35,71	52,63	21,43	52,63	14,29
3	57,89	78,57	63,16	21,43	42,11	14,29
4	52,63	64,29	42,11	42,86	42,11	100
5	26,32	42,86	47,37	100	36,84	100
6	31,58	28,57	57,89	14,29	57,89	14,29
7	47,37	50	47,37	100	47,37	14,29
8	47,37	21,43	36,84	28,57	36,84	100
9	57,89	21,43	26,32	71,43	31,58	100
10	63,16	42,86	36,84	100	47,37	0
11	47,37	71,43	36,84	28,57	52,63	7,143
12	52,63	64,29	42,11	42,86	42,11	100
13	42,11	100	63,16	35,71	31,58	100
14	63,16	50	57,89	21,43	42,11	100
15	42,11	64,29	57,89	71,43	52,63	0
16	47,37	7,143	31,58	100	52,63	21,43
17	31,58	85,71	36,84	28,57	42,11	21,43
18	52,63	64,29	42,11	64,29	52,63	14,29
19	57,89	42,86	52,63	64,29	36,84	21,43
20	36,84	21,43	31,58	64,29	52,63	42,86
Mean accuracy	48,16	50,36	45,79	53,21	45,26	45,36
Standard deviation	10,3	23,42	10,8	28,97	7,699	40,99
Sensitivity	0.507	-	0.485	-	0.479	-
Specificity	0.455	-	0.431	-	0.427	-
PPV	0.485	-	0.46	-	0.445	-
NPV	0.448	-	0.456	-	0.461	-

7.3.1.8 The correlation with regressed data classifier

Description

The results of the correlation with regressed data classifier are shown in table 7.8 and figure 7.10. The best classification accuracy is obtained when the features were calculated using the third atlas level (234 parcels). Both the mean accuracy and the compactness of both the validation and the optional validation set result distribution increase consistently when more features are used. The highest accuracy is reached when all 20 features are used (mean accuracy = $\pm$83%) The compactness when all 20 features are used is extremely high (SD = 6.359 and 8.268 for the validation and optional validation set respectively), reflecting the very high quality and consistency of the feature set. Three positive outliers are present: the seventh iteration when all features are used (average accuracy = $\pm$90.9%, sensitivity = 0.9, specificity = 0.889, PPV = 0.9, NPV = 0.889), the fourteenth iteration when all features are used (average accuracy = $\pm$90.9%, sensitivity = 0.9, specificity = 0.889, PPV = 0.9, NPV = 0.889) and the twentieth iteration when 10 features are used (average accuracy = $\pm$88.5%, sensitivity = 0.889, specificity = 0.8, PPV = 0.8, NPV = 0.889). This feature set contains the only iteration where an accuracy of 100% is reached for the validation set (first iteration).

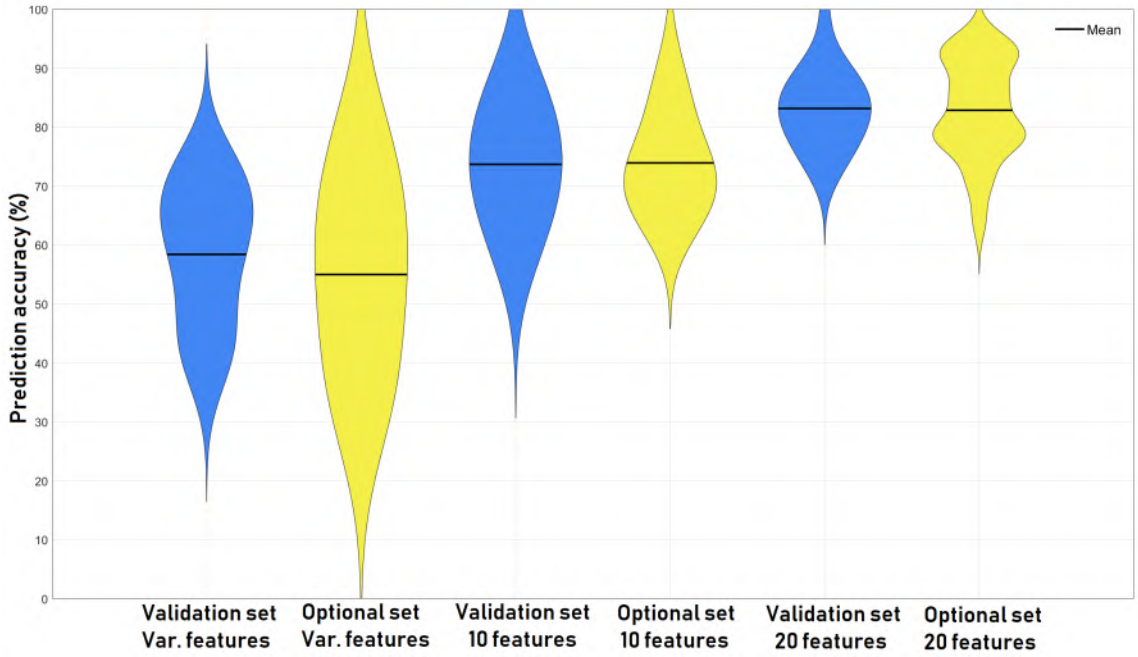


Figure 7.10: Violin plot of the results of the correlation with regressed data classifier.

Table 7.8: Best results of the correlation with regressed data classifier (Atlas3).

Iteration	Var. features		10 features		20 features	
	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)
1	68,42	64,29	63,16	71,43	100	78,57
2	57,89	78,57	63,16	85,71	84,21	78,57
3	57,89	35,71	68,42	71,43	84,21	78,57
4	57,89	35,71	84,21	64,29	89,47	85,71
5	73,68	50	73,68	64,29	78,95	92,86
6	42,11	50	84,21	85,71	78,95	78,57
7	73,68	50	63,16	78,57	89,47	92,86
8	63,16	57,14	84,21	71,43	84,21	92,86
9	63,16	71,43	57,89	64,29	84,21	71,43
10	73,68	35,71	73,68	71,43	73,68	71,43
11	42,11	35,71	89,47	71,43	78,95	78,57
12	63,16	64,29	73,68	78,57	78,95	85,71
13	68,42	35,71	73,68	71,43	84,21	85,71
14	42,11	78,57	57,89	71,43	89,47	92,86
15	52,63	64,29	84,21	71,43	73,68	92,86
16	36,84	57,14	73,68	64,29	84,21	78,57
17	47,37	28,57	78,95	85,71	84,21	64,29
18	68,42	71,43	73,68	64,29	73,68	85,71
19	47,37	78,57	68,42	78,57	78,95	92,86
20	68,42	57,14	84,21	92,86	89,47	78,57
Mean accuracy	58,42	55	73,68	73,93	83,16	82,86
Standard deviation	11,64	15,99	9,267	8,23	6,359	8,268
Sensitivity	0.611	-	0.706	-	0.858	-
Specificity	0.558	-	0.771	-	0.807	-
PPV	0.58	-	0.761	-	0.815	-
NPV	0.589	-	0.715	-	0.85	-

7.3.1.9 The mutual information with regressed data classifier

Description

The results of the mutual information with regressed data classifier are shown in table 7.9 and figure 7.11. The best classification accuracy is obtained when the features were calculated using the third atlas level (234 parcels). The mean accuracy as well as the compactness for both the validation an optional validation set result distributions remain fairly similar when more features are used. The highest accuracy is obtained when all features are used (mean accuracy = ±79%). Drawing conclusions from this fact is difficult. It is possible (this is not definitive) that the variable amount of features was high (between 10 and 20), resulting in similar results. The small increase in mean accuracy and compactness when all 20 features are used could be explained by the small increase in used features. Three positive outliers are present: the second iteration when 10 features are used (average accuracy = ±90.9%), the sixth iteration when a variable amount of features is used (average accuracy = ±87.6%, sensitivity = 0.889, specificity = 0.9, PPV = 0.889, NPV = 0.9) and the ninth iteration when a variable amount of features is used (average accuracy = ±87.6%, sensitivity = 0.9, specificity =). The presence of two outliers in the variable amount of features column strengthens the suspicion that the variable amount of features that is used was high.

Table 7.9: Best results of the mutual information with regressed data classifier (Atlas3).

Iteration	Var. features		10 features		20 features	
	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)	Val.(%)	Opt.(%)
1	78,95	64,29	68,42	78,57	89,47	71,43
2	78,95	57,14	89,47	92,86	73,68	85,71
3	63,16	50	73,68	85,71	84,21	64,29
4	52,63	71,43	73,68	71,43	78,95	78,57
5	78,95	92,86	78,95	92,86	84,21	92,86
6	89,47	85,71	78,95	71,43	89,47	64,29
7	68,42	100	78,95	71,43	84,21	78,57
8	73,68	64,29	68,42	78,57	89,47	78,57
9	89,47	85,71	84,21	64,29	68,42	85,71
10	73,68	71,43	84,21	71,43	73,68	78,57
11	78,95	78,57	84,21	85,71	78,95	92,86
12	68,42	50	68,42	78,57	78,95	78,57
13	78,95	85,71	68,42	71,43	78,95	85,71
14	73,68	85,71	63,16	57,14	78,95	64,29
15	73,68	78,57	84,21	71,43	73,68	85,71
16	63,16	78,57	63,16	64,29	73,68	78,57
17	78,95	71,43	63,16	85,71	78,95	71,43
18	84,21	85,71	78,95	64,29	68,42	78,57
19	73,68	92,86	73,68	85,71	84,21	78,57
20	73,68	71,43	78,95	71,43	73,68	92,86
Mean accuracy	74,74	76,07	75,26	75,71	79,21	79,29
Standard deviation	8,584	13,6	7,824	9,689	6,332	8,719
Sensitivity	0.772	-	0.779	-	0.768	-
Specificity	0.724	-	0.727	-	0.817	-
PPV	0.735	-	0.74	-	0.805	-
NPV	0.761	-	0.767	-	0.78	-

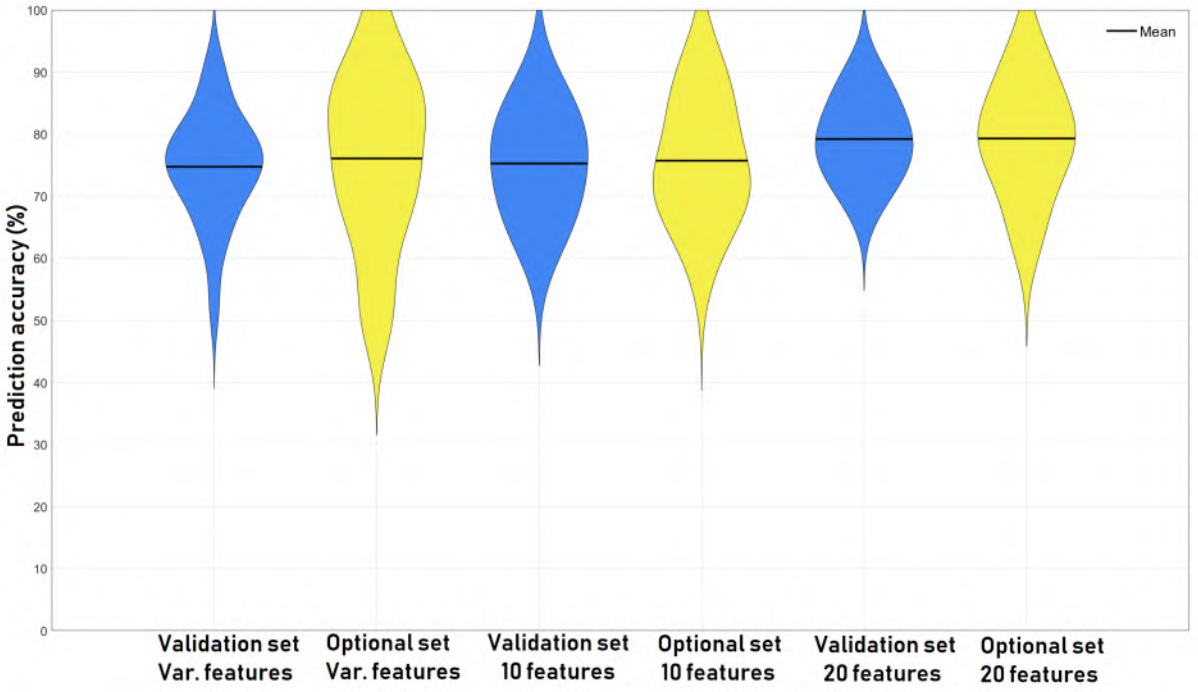


Figure 7.11: Violin plot of the results of the mutual information with regressed data classifier.

7.3.2 Combined-feature classifiers

7.3.2.1 The structural feature classifier

Description

The structural feature classifier is obtained by combining all three structural subtype feature sets. The results of the structural feature classifier are shown in table 7.10 and figure 7.12. The mean accuracy differs significantly between the validation and optional validation set (mean accuracy = 58.68% for the validation set, mean accuracy = 64.29% for the optional validation set). The compactness is similar for both validation sets.

Table 7.10: Best results of the structural feature classifier

Iteration	36 features	
	Val.(%)	Opt.(%)
1	42,11	71,43
2	47,37	42,86
3	57,89	71,43
4	63,16	28,57
5	52,63	57,14
6	52,63	64,29
7	84,21	64,29
8	57,89	71,43
9	68,42	71,43
10	57,89	57,14
11	57,89	78,57
12	63,16	42,86
13	52,63	71,43
14	57,89	50

15	78,95	71,43
16	63,16	78,57
17	42,11	64,29
18	52,63	85,71
19	57,89	78,57
20	63,16	64,29
Mean accuracy	58,68	64,29
Standard deviation	10,43	14,29
Sensitivity	0.613	-
Specificity	0.561	-
PPV	0.58	-
NPV	0.594	-

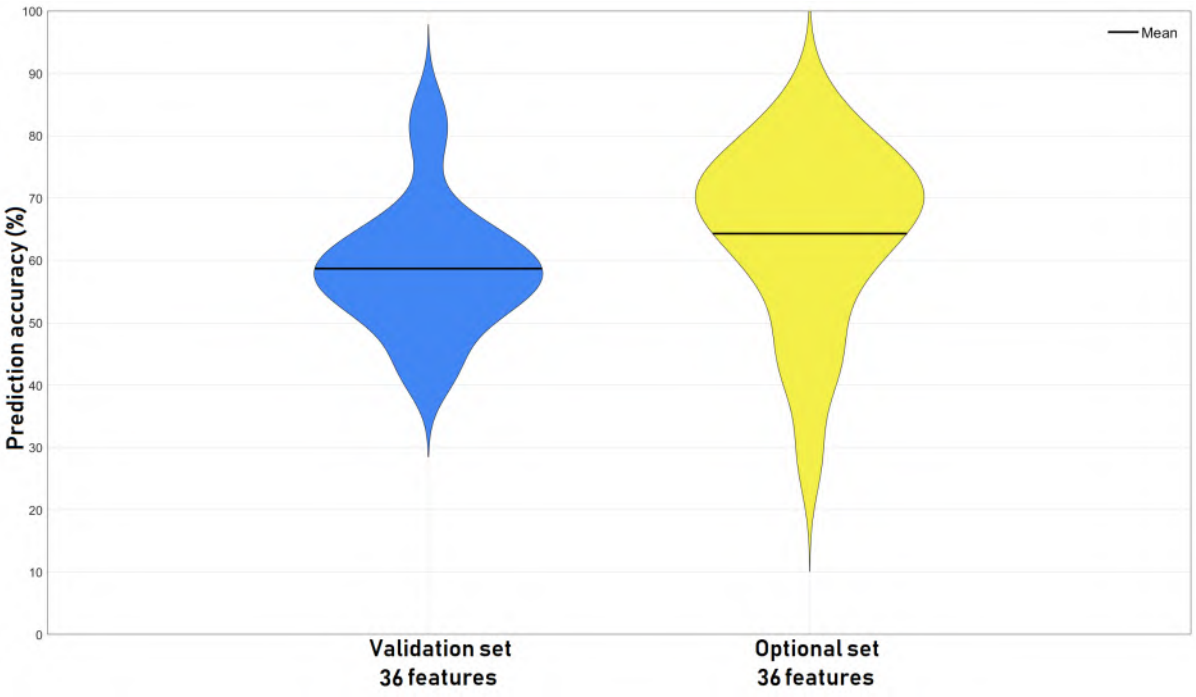


Figure 7.12: Violin plot of the results of the structural feature classifier.

7.3.2.2 The intensity feature classifier

Description

The intensity feature classifier is obtained by combining the absolute and relative intensity feature sets. The results of the intensity feature classifier are shown in table 7.11 and figure 7.13. Both mean accuracy and the compactness are similar between the validation and optional validation set. The mean accuracy is ±70%. Two positive outliers are present: the second iteration (average accuracy = ±80%, sensitvity = 0.89, specificity = 0.8, ppv = 0.8, npv = 0.89) and the seventeenth iteration (average accuracy = ±89.5%, sensitvity = 0.8, specificity = 0.78, ppv = 0.8, npv = 0.78).

Table 7.11: Best results of the intensity feature classifier

Iteration	40 features	
	Val.(%)	Opt.(%)
1	84,21053	64,28571
2	84,21053	78,57143
3	68,42105	64,28571
4	78,94737	50
5	63,15789	57,14286
6	78,94737	71,42857
7	68,42105	85,71429
8	63,15789	78,57143
9	73,68421	78,57143
10	89,47368	71,42857
11	84,21053	64,28571
12	63,15789	85,71429
13	73,68421	57,14286
14	63,15789	64,28571
15	63,15789	78,57143
16	73,68421	71,42857
17	78,94737	100
18	57,89474	71,42857
19	31,57895	28,57143
20	78,94737	71,42857
Mean accuracy	71,05263	69,64286
Standard deviation	12,94847	14,99642
Sensitivity	0.738	-
Specificity	0.684	-
PPV	0.695	-
NPV	0.728	-

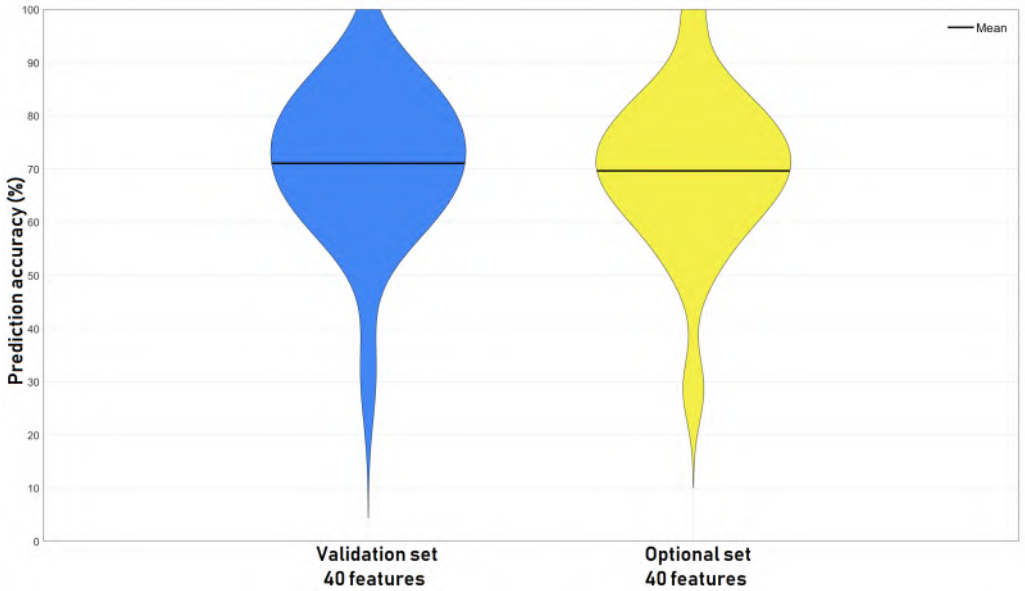


Figure 7.13: Violin plot of the results of the intensity feature classifier.

7.3.2.3 The connectivity feature classifier

Description

The connectivity feature classifier is obtained by combining the correlation and mutual information with regressed data set features. The results of the connectivity feature classifier is shown in table 7.12 and figure 7.14. The results of the validation set (both mean accuracy and SD) are better compared to the results of the optional validation set; the increase in standard deviation is the most significant (from 8.95 to 18.38). One positive outlier is present: the fourteenth iteration (average accuracy = ±73.5%, sensitivity = 0.7, specificity = 0.67, ppv = 0.7, npv = 0.67).

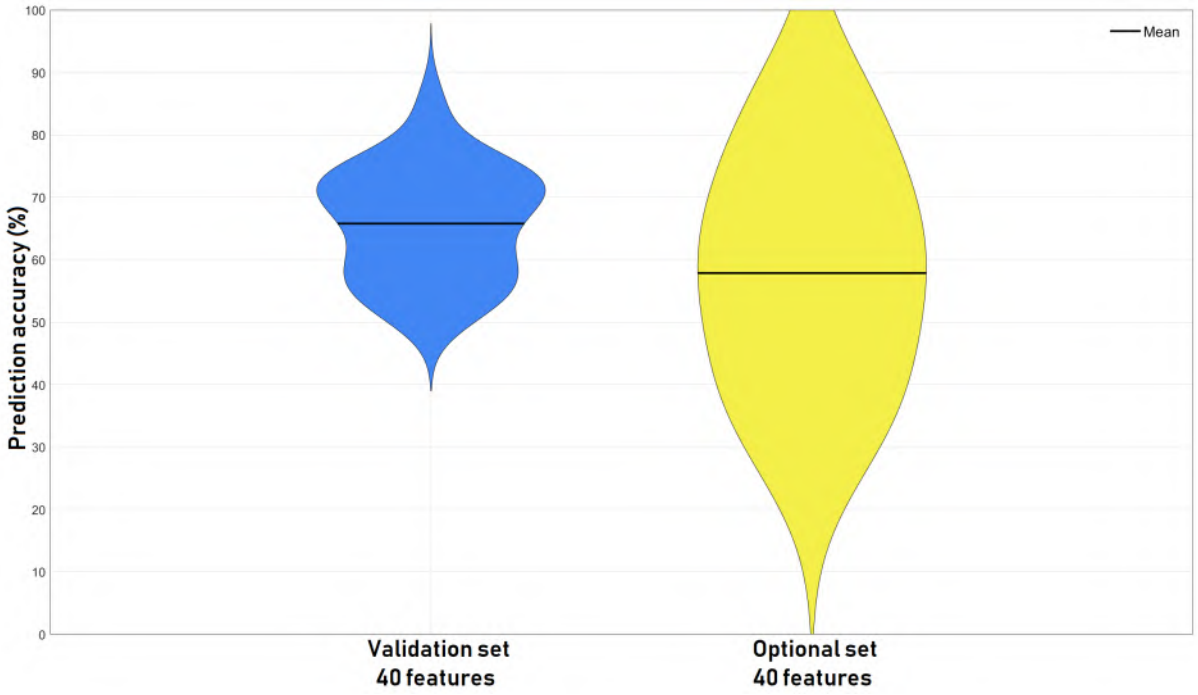


Figure 7.14: Violin plot of the results of the connectivity feature classifier.

Table 7.12: Best results of the connectivity feature classifier

Iteration	40 features	
	Val.(%)	Opt.(%)
1	68,42	57,14
2	84,21	64,29
3	73,68	64,29
4	52,63	50
5	57,89	35,71
6	68,42	71,43
7	52,63	50
8	73,68	85,71
9	68,42	71,43
10	57,89	35,71
11	52,63	85,71
12	73,68	35,71
13	73,68	42,86
14	68,42	78,57
15	63,16	28,57

16	73,68	64,29
17	63,16	57,14
18	57,89	85,71
19	73,68	57,14
20	57,89	35,71
Mean accuracy	65,79	57,86
Standard deviation	8,955	18,38
Sensitivity	0.688	-
Specificity	0.629	-
PPV	0.645	-
NPV	0.672	-

7.3.2.4 The intensity and connectivity feature classifier

Description

The intensity and connectivity feature classifier is obtained by combining four feature sets: the absolute intensity (atlas3) feature set, the relative intensity (atlas5) feature set, the correlation with regressed data (atlas3) feature set and the mutual information with regressed data (atlas3) feature set. The results of the intensity and connectivity feature classifier is shown in figure 7.13 and figure 7.15. This classifier has the best performance of all trained classifiers. It has a mean accuracy of ±88.7% and a mean standard deviation of ±6.97. The compactness of both result distributions is very high, showing the high quality of the feature set. Multiple positive outliers are present: the sixth iteration (average accuracy = ±90.2%, sensitivity = 1, specificity = 0.9, ppv = 0.9, nvp = 1), the first, fourth, seventh, eighth, thirteenth, sixteenth and seventeenth iteration (average accuracy = ±92.1%, sensitivity = 0.89, specificity = 0.8, ppv = 0.8, npv = 0.89) and the nineteenth iteration (average accuracy = ±94.7%, sensitivity = 0.9, specificity = 0.89, ppv = 0.9, npv = 0.89).

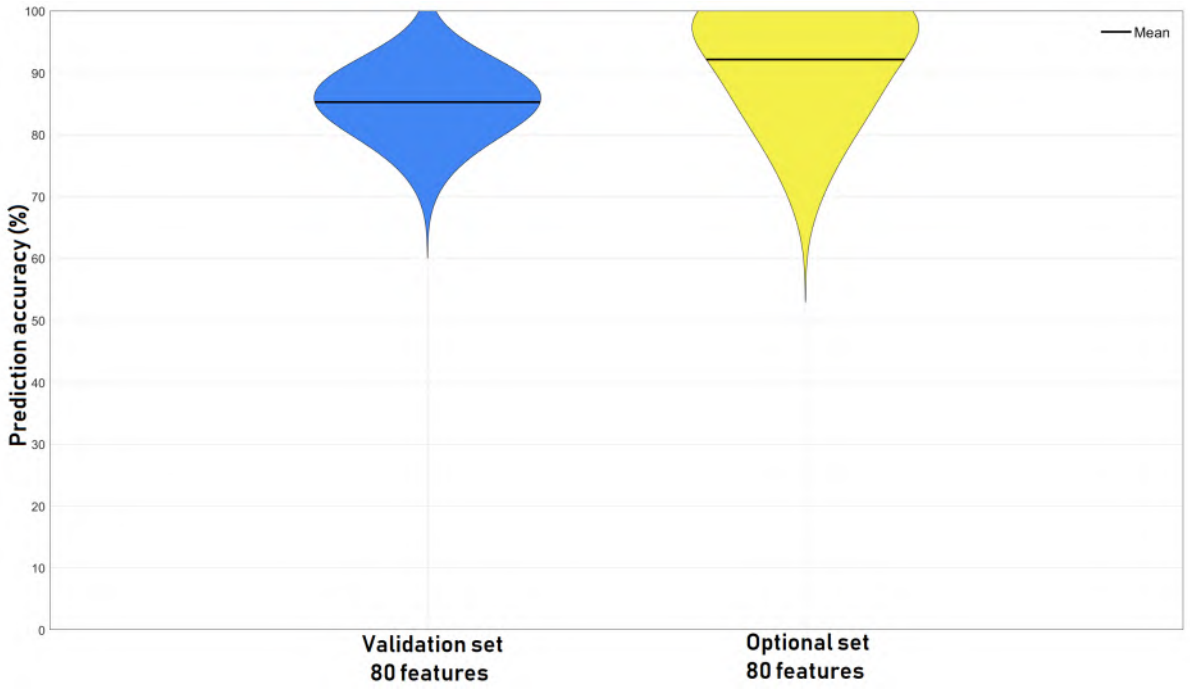


Figure 7.15: Violin plot of the results of the connectivity feature classifier.

Table 7.13: Best results of the intensity and connectivity feature classifier.

Iteration	80 features	
	Val.	Opt.
1	84,21053	100
2	78,94737	100
3	78,94737	85,71429
4	84,21053	100
5	84,21053	71,42857
6	94,73684	85,71429
7	84,21053	100
8	84,21053	100
9	89,47368	92,85714
10	89,47368	78,57143
11	89,47368	92,85714
12	89,47368	92,85714
13	84,21053	100
14	89,47368	85,71429
15	89,47368	92,85714
16	84,21053	100
17	84,21053	100
18	78,94737	85,71429
19	89,47368	100
20	73,68421	78,57143
Mean accuracy	85,26316	92,14286
Standard deviation	5,007648	8,945473
Sensitivity	0.879	-
Specificity	0.827	-
PPV	0.835	-
NPV	0.872	-

Chapter 8

Discussion

This chapter contains the discussion of the results shown and described in chapter 7. The results of each feature subtype will be discussed with respect to other feature subtypes belonging to the same feature type (intensity, connectivity and structural). Focus will be given to accuracy differences, differences in compactness of result distributions and changes related to the amount of used features. The results of each feature type will also be discussed with respect to the other feature types, similar focus as in the previous part is used. The combined feature classifiers are discussed afterwards. Finally the influence of the atlas level will be discussed.

8.1 Part 1: Feature type specific

8.1.1 Intensity features

No significant difference in accuracy exists between the absolute and relative intensity feature classifiers. The accuracy of both classifiers is positively correlated with the amount of features that is used and an increase of compactness is noticed when all 20 features are used. From a result-based point of view, no clear preference exists between both feature subtypes (contrary to the clinical relevance point of view, see section 5.3.3).

8.1.2 Connectivity features

Two large differences are noticeable for the connectivity subtype classifiers: the use of a regressed versus non-regressed data set and the change in accuracy and compactness between correlation-based and mutual information-based classifiers with respect to the used data set.

The first noticeable difference is a significant increase in accuracy for both the correlation-based and mutual information-based classifiers when the regressed data set is used. The increase in accuracy is however much bigger for the mutual information-based classifiers (an increase of $\pm$35% versus an increase of $\pm$6%). A considerable increase in compactness is also noticeable for each feature subtype classifier when the regressed data set is used. The added value of global signal regression (see section 5.4.2.1) is clear. The possibility of the introduction of anti-correlated connectivities (see [65]) leading to false conclusions is partly negated by the fact that the difference between two connectivity values is calculated. An increase or decrease of both values would not change the difference value between them significantly.

The second noticeable difference is the fact that the correlation-based classifiers are able to achieve adequate accuracy results (mean accuracy = $\pm$77.5% when all 20 features are used) when using the

non-regressed data set while the mutual information-based classifiers are not able to do this (mean accuracy = ±48% when all 20 features are used). The reason why the mutual information feature set is not able to correctly distinguish between the healthy controls and the depression group is counter-intuitive. Generally speaking, high variability of feature values between different people both in the healthy control group and the depression group is the main reason for the limited achieved accuracies (this is most noticeable in the structural feature sets, see section 8.2). A lack of variation however is the main source of inadequate features when mutual information is used with the non-regressed data set. The range of the group matrices (obtained in the group averaging step in the feature selection process, see section 5.4.2.4) is [4.2542,4.6980] for the healthy controls and [4.2914,4.8101] for the depression group. The difference matrix obtained has a range of [-0.2316,0.1158]. This range is too low, which results in inadequate features. Mutual information has a possible range of $[0,+\infty[$ and the small range of values in the group matrices could indicate the presence of a strong global signal. The high increase in accuracy when the regressed data set is used strengthens this suspicion.

8.1.3 Structural features

Not much is to be said about the comparison of the structural feature subtypes from a result-based point of view. All feature subtypes are inadequate for consistently predicting depression. The reason for this is given in section 8.2.

8.2 Part 2: Feature type comparison

The best performing classifiers are trained with the connectivity features (only when the regressed data set is used, see section 8.1.2). An average accuracy of ±83% is obtained with correlation and ±79% is obtained with mutual information; both feature sets have a positive outlier with an accuracy of ±90.9%. The second best performing classifiers are trained with the intensity features. An average accuracy of ±75% is obtained with absolute intensity and an average accuracy of ±73% is obtained with relative intensity; both feature sets have a positive outlier with an accuracy of ±87% (absolute intensity) and ±88% (relative intensity). The worst performing classifiers are trained with the structural features. An average accuracy of ±59.5% is obtained with left hemisphere thickness, an average accuracy of ±56% is obtained with right hemisphere thickness and an average accuracy of ±61% is obtained with parcel volume. No considerable outliers, when compared to the other feature type classifiers, are present.

A significant difference in accuracy is noticeable between the three different feature types. A small (±5% to ±10%) decrease in accuracy is noticed between the connectivity and intensity features. A possible explanation for this phenomenon could be the fact that both features are calculated from the same data (fMRI data), but that more information is lost during the feature selection process of the intensity features. Intensity features are calculated by averaging the time series, thus ignoring the variation of the brain activity through time. Connectivity features do not ignore this information, possibly leading to better feature sets. The higher quality of the connectivity feature sets is also confirmed by the higher compactness of the result distributions (SD of ±6.3 versus ±9.8 for the validation set and SD of ±8.5 versus ±12 for the optional validation set).

Structural features perform worse than both intensity and connectivity features. Decreases in accuracy of ±20% to ±30% are present when compared to the two other feature types. Similar decreases in compactness (SD of ±13 versus ±6.3 or ±9.8 for the validation set and SD of ±20% versus ±8.5 or ±12 for the optional validation set) are present. A possible explanation for the large decrease in both accuracy and compactness is the fact that structural features reflect anatomical changes while intensity and connectivity reflect functional changes. Changes in cortical thickness or volumes would

reflect possible atrophy in the brain. Measurable brain atrophy could only be the result of very severe forms of depression where a patient experienced multiple depressive episodes. A simplification has been used in this master's dissertation: differences in severity and duration of depression between patients is ignored. This simplification is needed to reduce the problem to a binary classification problem. Statistically significant structural features have been found (see secton 4), showing the presence of anatomical changes. The rate of change (due to possible atrophy) however is not constant within the depression group as some people within the depression group have had more depressive episodes/suffer longer from depression than others. If more information about the duration and amount of depressive episodes is known (as well as other factors that could influence anatomical changes such as used medications or therapies) structural features could probably be used to better predict both depression and depression severity. This problem could likely be solved using regression machine learning techniques.

8.3 Part 3: Combined feature classifiers

Combined feature classifiers are investigated as the inclusion of more features could possibly lead to an increase in accuracy. A first step in the creation of combined feature classifiers is the combination of the different feature subtypes. Three combined feature classifiers are defined: an intensity feature classifier, a connectivity feature classifier and a structural feature classifier. The performance of these classifiers will be compared to the performance of the best subtype classifiers. All combined feature classifiers are trained once with all combined features. Due to the decrease in classification performance compared to the subtype feature classifiers, no final combined feature containing all defined features is trained and tested.

8.3.1 The intensity feature classifier

The performance of the intensity feature classifier is slightly lower than both the absolute and relative intensity feature classifier ($\pm$70% compared to $\pm$74.5% and $\pm$73% respectively). A possible explanation is the fact that by introducing more features also more noise is introduced. The increased noise makes it more difficult to correctly distinguish healthy controls and depression patients.

8.3.2 The connectivity feature classifier

The performance of the connectivity feature classifier is much lower than the connectivity subtype classifiers (except from the mutual information on the non-regressed data classifier). This is counter-intuitive as the connectivity subtype feature sets are the best performing feature sets, both with very high compactness, and because the intensity feature classifier did not share a similar drop in accuracy and compactness. No clear explanation for this phenomenon can be given.

8.3.3 The structural feature classifier

Contrary to the other combined classifiers, the mean accuracy and compactness is not lower than the feature subtype classifiers, but is even increased slightly. The increase however is too small to be considered significant.

8.3.4 The intensity and connectivity feature classifier

The intensity and conncectivity feature classifier has the best performance of all classifiers. This is both intuitive and counter-intuitive. The addition of more features to distinguish different classes

normally leads to higher classification accuracies, so the aggregation of all high accuracy feature sets into a single feature set intuitively leads to higher classification results. This is however also counter-intuitive as the other combined feature classifiers built from the feature subtypes that are included in this feature set (the intensity and connectivity feature classifier, see section 8.3.1 and 8.3.2) have lower accuracies than the feature subtype classifiers.

This classifier performs very well, but caution should be used when forming conclusions. 80 features are present in this feature set and only 73 people are used for training (see section 6.1.3). SVMs, and kernel machines in general, are capable of classification when more features than different training instances are present. The use of more features however results in a higher chance of overfitting. To confirm the ability of this classifier more validation is needed.

8.4 Part 4: Performance with respect to atlas level

A final aspect of the results that is noticed is the fact that four of the six feature subtypes that are calculated with the Lausanne brain atlas (see section 5.2.1) have the best results when atlas level three is used (234 parcels). This is shown in figure 8.1 where the mean accuracy of both the validation and optional validation set (when all 20 features are used) of the absolute and relative intensity as well as the correlation with regressed data feature set are shown. The average accuracy of each atlas level is denoted by the thicker brown line. This figure shows that on average the best accuracy is obtained at atlas level 3 and that the higher atlas levels (atlas level 4 and 5) have a slightly higher accuracy than the lower atles level (atlas 1 and 2). This shows that a parcellation that distinguishes more brain parcels is advantageous for feature calculation.

Atlas level one defines parcels that usually contain multiple unique brain regions; features calculated from these parcels will be the average of several brain regions and therefore will not contain all information from these brain regions. Atlas level five defines brain regions that are very small. Slight positioning variations between different scans due to the preprocessing process have considerable influence on the feature values, reducing the consistent differences needed for high quality features. The parcels defined by atlas level three circumvent both problems. A possible explanation for this phenomenon, aside from coincidence, is that the parcel size defined in atlas level three is an optimal middle ground. The viability of this hypothesis is strengthened when the absolute intensity features and mutual information with regressed data features are discussed (see section 5.3.3.2 and 5.4.3.2). Both feature sets contain several subdivisions of a single brain region (both use atlas level 3), showing that the subdivision leads to higher classification accuracies, but does not use the highest subdivision (atlas level 5) available as the classification accuracies are again reduced.

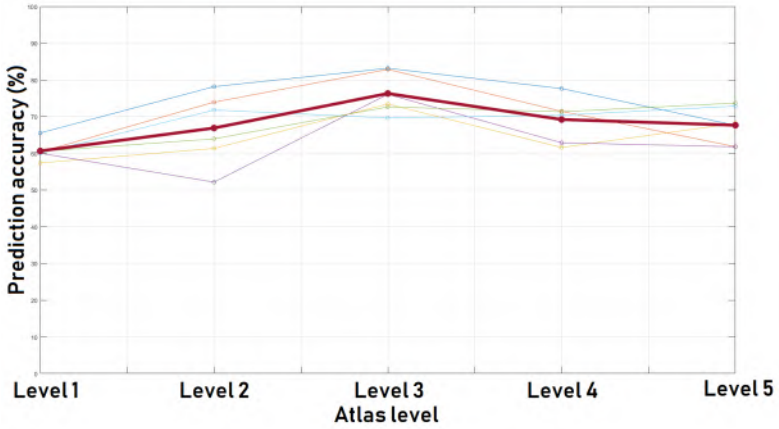


Figure 8.1: Violin plot of the results of the connectivity feature classifier.

Chapter 9

Conclusion

In this master's dissertation a computer-aided diagnosis tool/classifier is built that is capable of diagnosing depression based on an fMRI scan. A data set of 106 people, 60 healthy controls and 46 people diagnosed with depression, was used to obtain features and train classifiers.

Three different feature types for classification were explored: intensity features, reflecting the average activity in the brain; connectivity features, reflecting the functional connectivity between brain regions and structural features, reflecting the thickness and volume of the brain. An important principle that was used in the selection process of the different feature types is the possibility of a simple interpretation: each feature can easily be linked to the corresponding brain regions and the clinical relevance of each feature can easily be determined by a physician.

Each feature type in itself consisted of several subtypes, a total of nine feature subtypes were defined and used to train the classifier. Two intensity feature subtypes were defined: absolute and relative intensity features. Two different connectivity measures were used to calculate functional connectivity between different brain regions: correlation and mutual information. An extra preprocessing step, global signal regression, was used resulting in two different data sets. Both connectivity measures were calculated on both data sets, resulting in four different connectivity feature subtypes. Three different structural feature subtypes were used: left and right cortical thickness and volume. Support vector machines were trained multiple times for each feature type resulting in a result distribution for each feature type and subtype. The properties of the result distributions were analyzed to interpret the viability of each feature subtype as a distinguishing factor for depression.

Every feature type and subtype did have at least some features that can be linked to depression and thus have some clinical relevance; some feature subtypes (absolute and relative intensity, left hemishpere cortical thickness and parcel volume) even have high clinical relevance as most of the features are closely linked with depression. This shows that, even when no prior assumptions about the disease have been made, the found features reflect brain anatomy and activity that are also found when no data-driven approach is used. The presence of clinical relevant features increases the diagnositc value of the found classifiers.

The best performing feature type is the connectivity feature type. An average accuracy of $\pm$83% is reached (validation set 83.16%, optional validation set 82,86%) when correlation is used on the global signal regressed data set, an average accuracy of $\pm$79% is reached (validation set 79.21%, optional validation set 82,86%) when mutual information is used. The highest performing trained classifier with correlation that is obtained has an average accuracy of 90.9% (validation set 89.47%, optional validation set 92.86%), this accuracy is reached twice in twenty training cycles. The highest performing trained classifier with mutual information that is obtained has an average accuracy of 87.7% (validation set

84.21%, optional validation set 92.86%). This shows that functional connectivity changes are highly reliable features that can be used for a possible computer-aided diagnosis tool for depression.

Intensity features also show some capability of correctly predicting depression with absolute intensity feature classifiers having an average accuracy of $\pm$75% and relative intensity features having an average accuracy of $\pm$73%. The best performing classifier has an average accuracy of 87.3% (validation set 89.47%, optional validation set 85.71%) for absolute intensity features.

Differences in structural features between healthy controls and patients with depression are proven to be statistically significant, but they are not capable of accurately predicting depression. Accuracies between 50% and 60% are obtained using structural features, showing the unreliability of this type of feature for classification.

Combining multiple feature subtypes from a single feature type did not result in an increased accuracy. The final combination of the four best performing feature sets (absolute intensity, relative intensity, correlation with regressed data and mutual information with regressed data) however results in the best performing classifier. This feature set has an average accuracy of $\pm$88.7% and the best performing classifier has an accuracy of $\pm$94.7% (validation set 89.47%, optional validation set 100%). Caution should be used as the classifiers trained with this feature set could have been overfitted due to the high amount of features.

A final remark that needs to be made is the fact that the size of the data set, while big compared to similar research [36], is rather small. Further research using other and larger data sets is needed to assess the applicability of the results and conclusions obtained in this master's dissertation.

Nevertheless it can be concluded that a computer-aided diagnosis tool based on resting state fMRI data can be a reliant method of diagnosing depression. Tools like this could help mental health professionals with the diagnosis of depression and could even be used in advance of a diagnostic interview, decreasing the workload of mental health professionals and waiting periods for patients.

Future work

The obtained algorithm for feature selection and the resulting classification systems can be expanded in multiple ways. More feature types, such as frequency and network based features could be investigated. The algorithm could be used to obtain feature sets and classification systems for other diseases such as post traumatic stress disorder or schizophrenia. Other data sets could be used to further validate and refine the algorithm and classifiers. Task-related fMRI data could be used to build a classifier. Other imaging techniques such as EEG or fNIRS could be used together with fMRI to increase the accuracy of the obtained classifiers. Bigger data sets could be used so that more complex classification systems, such as random forests or artificial neural networks could be applied.

From a personal point of view the most important expansion would be the inclusion of the ability to not only diagnose depression, but also correctly predict specific subtypes of depression. The ability to diagnose subtypes of depression that are linked to efficacy of medication types and treatment options could decrease the trail period for the patient. Considerable research needs to be done to make this possible, but the potential benefit of alternative diagnosis tools in psychology and psychiatry is not to be understated.

Bibliography

[1] W. H. Organization *et al.*, "Depression and other common mental disorders: global health estimates," Tech. Rep. (World Health Organization, 2017).

[2] J. P. Hornak, "The basics of mri," *http://www. cis. rit. edu/htbooks/mri* (2006).

[3] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince, *MRI from Picture to Proton* (Cambridge university press, 2017).

[4] D. P. V. Schuerbeek, "Functional mri lecture 2," (2018).

[5] D. Saldanha, N. Kumar, V. Ryali, K. Srivastava, and A. Pawar, "Serum serotonin abnormality in depression," *Medical Journal Armed Forces India* **65**, 108 (2009).

[6] J. Hallas, "Evidence of depression provoked by cardiovascular medication: a prescription sequence symmetry analysis," *Epidemiology* , 478 (1996).

[7] H. Durdle, L. H. Lundahl, C.-E. Johanson, and M. Tancer, "Major depression: the relative contribution of gender, mdma, and cannabis use," *Depression and anxiety* **25**, 241 (2008).

[8] C. Heim, D. J. Newport, T. Mletzko, A. H. Miller, and C. B. Nemeroff, "The link between childhood trauma and depression: insights from hpa axis studies in humans," *Psychoneuroendocrinology* **33**, 693 (2008).

[9] S. R. Penzak, Y. S. Reddy, and S. R. Grimsley, "Depression in patients with hiv infection," *American Journal of Health-System Pharmacy* **57**, 376 (2000).

[10] J. A. Foster and K.-A. M. Neufeld, "Gut–brain axis: how the microbiome influences anxiety and depression," *Trends in neurosciences* **36**, 305 (2013).

[11] K. W. Kelley, J. C. O'Connor, M. A. Lawson, R. Dantzer, S. L. Rodriguez-Zas, and R. H. McCusker, "Aging leads to prolonged duration of inflammation-induced depression-like behavior caused by bacillus calmette-guerin," *Brain, behavior, and immunity* **32**, 63 (2013).

[12] L. A. Jelenchick, J. C. Eickhoff, and M. A. Moreno, ""facebook depression?" social networking site use and depression in older adolescents," *Journal of Adolescent Health* **52**, 128 (2013).

[13] D. G. Blazer, *The Age of Melancholy:" Major Depression" and its Social Origin* (Routledge, 2012).

[14] S. Eljamel and K. Slavin, *Neurostimulation: Principles and Practice* (John Wiley & Sons, 2013).

[15] D. Pagnin, V. de Queiroz, S. Pini, and G. B. Cassano, "Efficacy of ect in depression: a meta-analytic review," *The journal of ECT* **20**, 13 (2004).

[16] A. P. Association *et al.*, *Diagnostic and statistical manual of mental disorders (DSM-5®)* (American Psychiatric Pub, 2013).

[17] D. Draulans, "Vlaamse psychiaters slaan alarm: 'mensen plegen zelfmoord omdat ze niet de juiste zorg krijgen'," *Knack* (2019).

[18] P. Casteels, "Iedereen in therapie: 'de zoektocht naar een psycholoog zal altijd een lijdensweg zijn."," *Knack* , 40–44 (2019).

[19] M. B. First, R. L. Spitzer, M. Gibbon, and J. B. Williams, *User's guide for the Structured clinical interview for DSM-IV axis I disorders SCID-I: clinician version* (American Psychiatric Pub, 1997).

[20] J. Ventura, R. P. Liberman, M. F. Green, A. Shaner, and J. Mintz, "Training and quality assurance with the structured clinical interview for dsm-iv (scid-i/p)," *Psychiatry research* **79**, 163 (1998).

[21] K. Wittkampf, H. van Ravesteijn, K. Baas, H. van de Hoogen, A. Schene, P. Bindels, P. Lucassen, E. van de Lisdonk, and H. van Weert, "The accuracy of patient health questionnaire-9 in detecting depression and measuring depression severity in high-risk groups in primary care," *General hospital psychiatry* **31**, 451 (2009).

[22] A. Pettersson, K. B. Boström, P. Gustavsson, and L. Ekselius, "Which instruments to support diagnosis of depression have sufficient accuracy? a systematic review," *Nordic journal of psychiatry* **69**, 497 (2015).

[23] B. Burle, L. Spieser, C. Roger, L. Casini, T. Hasbroucq, and F. Vidal, "Spatial and temporal resolutions of eeg: Is it really black and white? a scalp current density view," *International Journal of Psychophysiology* **97**, 210 (2015).

[24] U. R. Acharya, V. K. Sudarshan, H. Adeli, J. Santhosh, J. E. Koh, S. D. Puthankatti, and A. Adeli, "A novel depression diagnosis index using nonlinear features in eeg signals," *European neurology* **74**, 79 (2015).

[25] O. Faust, P. C. A. Ang, S. D. Puthankattil, and P. K. Joseph, "Depression diagnosis support system based on eeg signal entropies," *Journal of mechanics in medicine and biology* **14**, 1450035 (2014).

[26] B. Hosseinifard, M. H. Moradi, and R. Rostami, in *2011 19th Iranian Conference on Electrical Engineering* (IEEE, 2011) pp. 1–4.

[27] W. Mumtaz, L. Xia, S. S. A. Ali, M. A. M. Yasin, M. Hussain, and A. S. Malik, "Electroencephalogram (eeg)-based computer-aided technique to diagnose major depressive disorder (mdd)," *Biomedical Signal Processing and Control* **31**, 108 (2017).

[28] Y. I. Sheline, D. M. Barch, J. L. Price, M. M. Rundle, S. N. Vaishnavi, A. Z. Snyder, M. A. Mintun, S. Wang, R. S. Coalson, and M. E. Raichle, "The default mode network and self-referential processes in depression," *Proceedings of the National Academy of Sciences* **106**, 1942 (2009).

[29] J. P. Hamilton, M. Farmer, P. Fogelman, and I. H. Gotlib, "Depressive rumination, the default-mode network, and the dark matter of clinical neuroscience," *Biological psychiatry* **78**, 224 (2015).

[30] X. Wang, Y. Ren, and W. Zhang, "Depression disorder classification of fmri data using sparse low-rank functional brain network and graph-based features," *Computational and mathematical methods in medicine* **2017** (2017).

[31] R. Bhaumik, L. M. Jenkins, J. R. Gowins, R. H. Jacobs, A. Barba, D. K. Bhaumik, and S. A. Langenecker, "Multivariate pattern analysis strategies in detection of remitted major depressive disorder using resting state functional connectivity," *NeuroImage: Clinical* **16**, 390 (2017).

[32] M. Wei, J. Qin, R. Yan, H. Li, Z. Yao, and Q. Lu, "Identifying major depressive disorder using hurst exponent of resting-state brain networks," *Psychiatry Research: Neuroimaging* **214**, 306 (2013).

[33] J. R. Sato, J. Moll, S. Green, J. F. Deakin, C. E. Thomaz, and R. Zahn, "Machine learning algorithm accurately detects fmri signature of vulnerability to major depression," *Psychiatry Research: Neuroimaging* **233**, 289 (2015).

[34] A. F. Marquand, J. Mourão-Miranda, M. J. Brammer, A. J. Cleare, and C. H. Fu, "Neuroanatomy of verbal working memory as a diagnostic biomarker for depression," *Neuroreport* **19**, 1507 (2008).

[35] C. H. Fu, J. Mourao-Miranda, S. G. Costafreda, A. Khanna, A. F. Marquand, S. C. Williams, and M. J. Brammer, "Pattern classification of sad facial processing: toward the development of neurobiological markers in depression," *Biological psychiatry* **63**, 656 (2008).

[36] M. J. Patel, A. Khalaf, and H. J. Aizenstein, "Studying depression using imaging and machine learning methods," *NeuroImage: Clinical* **10**, 115 (2016).

[37] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical parametric mapping: the analysis of functional brain images* (Elsevier, 2011).

[38] S. Whitfield-Gabrieli and A. Nieto-Castanon, "Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks," *Brain connectivity* **2**, 125 (2012).

[39] S. A. Huettel, A. W. Song, G. McCarthy, *et al.*, *Functional magnetic resonance imaging*, Vol. 1 (Sinauer Associates Sunderland, MA, 2004).

[40] G. B. Forbes, *Human body composition: growth, aging, nutrition, and activity* (Springer Science & Business Media, 2012).

[41] W. Greiner, *Quantum mechanics: an introduction* (Springer Science & Business Media, 2011).

[42] R. J. Gillies, I. Robey, and R. A. Gatenby, "Causes and consequences of increased glucose metabolism of cancers," *Journal of Nuclear Medicine* **49**, 24S (2008).

[43] G. Bush, J. A. Frazier, S. L. Rauch, L. J. Seidman, P. J. Whalen, M. A. Jenike, B. R. Rosen, and J. Biederman, "Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fmri and the counting stroop," *Biological psychiatry* **45**, 1542 (1999).

[44] E. Alpaydin, *Introduction to machine learning* (MIT press, 2009).

[45] R. Duprat, S. Desmyter, K. van Heeringen, D. Van den Abbeele, H. Tandt, J. Bakic, G. Pourtois, J. Dedoncker, M. Vervaet, S. Van Autreve, *et al.*, "Accelerated intermittent theta burst stimulation treatment in medication-resistant major depression: A fast road to remission?" *Journal of affective disorders* **200**, 6 (2016).

[46] M. Brant-Zawadzki, G. D. Gillan, and W. R. Nitz, "Mp rage: a three-dimensional, t1-weighted, gradient-echo sequence–initial experience in the brain." *Radiology* **182**, 769 (1992).

[47] J. Ashburner, G. Barnes, C. Chen, J. Daunizeau, G. Flandin, K. Friston, S. Kiebel, J. Kilner, V. Litvak, R. Moran, *et al.*, "Spm12 manual," *URL: http://www. fil. ion. ucl. ac. uk/spm/doc/spm12 manual. pdf* (2016).

[48] L. Yan, Y. Zhuo, Y. Ye, S. X. Xie, J. An, G. K. Aguirre, and J. Wang, "Physiological origin of low-frequency drift in blood oxygen level dependent (bold) functional magnetic resonance imaging (fmri)," *Magnetic resonance in medicine* **61**, 819 (2009).

[49] A. C. Evans, D. L. Collins, S. Mills, E. Brown, R. Kelly, and T. M. Peters, in *1993 IEEE conference record nuclear science symposium and medical imaging conference* (IEEE, 1993) pp. 1813–1817.

[50] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the structural core of human cerebral cortex," *PLoS biology* **6**, e159 (2008).

[51] J. Akiyoshi, K. Hieda, Y. Aoki, and H. Nagayama, "Frontal brain hypoactivity as a biological substrate of anxiety in patients with panic disorders," *Neuropsychobiology* **47**, 165 (2003).

[52] D. Hecht, "Depression and the hyperactive right-hemisphere," *Neuroscience research* **68**, 77 (2010).

[53] J. B. Henriques and R. J. Davidson, "Left frontal hypoactivation in depression." *Journal of abnormal psychology* **100**, 535 (1991).

[54] M. S. George, T. A. Ketter, and R. M. Post, "Prefrontal cortex dysfunction in clinical depression," *Depression* **2**, 59 (1994).

[55] W. C. Drevets, J. L. Price, J. R. Simpson Jr, R. D. Todd, T. Reich, M. Vannier, and M. E. Raichle, "Subgenual prefrontal cortex abnormalities in mood disorders," *Nature* **386**, 824 (1997).

[56] L. R. Baxter, J. M. Schwartz, M. E. Phelps, J. C. Mazziotta, B. H. Guze, C. E. Selin, R. H. Gerner, and R. M. Sumida, "Reduction of prefrontal cortex glucose metabolism common to three types of depression," *Archives of general psychiatry* **46**, 243 (1989).

[57] T. Paus and J. Barrett, "Transcranial magnetic stimulation (tms) of the human frontal cortex: implications for repetitive tms treatment of depression," *Journal of Psychiatry and Neuroscience* **29**, 268 (2004).

[58] Z. Nahas, C. C. Teneback, A. Kozel, A. M. Speer, C. DeBrux, M. Molloy, L. Stallings, K. M. Spicer, G. Arana, D. E. Bohning, *et al.*, "Brain effects of tms delivered over prefrontal cortex in depressed adults: role of stimulation frequency and coil–cortex distance," *The Journal of neuropsychiatry and clinical neurosciences* **13**, 459 (2001).

[59] M. P. Van Den Heuvel and H. E. H. Pol, "Exploring the brain network: a review on resting-state fmri functional connectivity," *European neuropsychopharmacology* **20**, 519 (2010).

[60] M. G. Kendall, "Rank correlation methods." (1948).

[61] J. Hlinka, M. Paluš, M. Vejmelka, D. Mantini, and M. Corbetta, "Functional connectivity in resting-state fmri: is linear correlation sufficient?" *Neuroimage* **54**, 2218 (2011).

[62] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, 2012).

[63] A. E. Desjardins, K. A. Kiehl, and P. F. Liddle, "Removal of confounding effects of global signal in functional mri analyses," *Neuroimage* **13**, 751 (2001).

[64] P. M. Macey, K. E. Macey, R. Kumar, and R. M. Harper, "A method for removal of global effects from fmri time series," *Neuroimage* **22**, 360 (2004).

[65] K. Murphy, R. M. Birn, D. A. Handwerker, T. B. Jones, and P. A. Bandettini, "The impact of global signal regression on resting state correlations: are anti-correlated networks introduced?" *Neuroimage* **44**, 893 (2009).

[66] Z. S. Saad, S. J. Gotts, K. Murphy, G. Chen, H. J. Jo, A. Martin, and R. W. Cox, "Trouble at rest: how correlation patterns and group differences become distorted after global signal regression," *Brain connectivity* **2**, 25 (2012).

[67] M. D. Fox, D. Zhang, A. Z. Snyder, and M. E. Raichle, "The global signal and observed anticorrelated resting state brain networks," *Journal of neurophysiology* **101**, 3270 (2009).

[68] I. Jolliffe, *Principal component analysis* (Springer, 2011).

[69] S. G. Disner, C. G. Beevers, E. A. Haigh, and A. T. Beck, "Neural mechanisms of the cognitive model of depression," *Nature Reviews Neuroscience* **12**, 467 (2011).

[70] M. D. Greicius, B. H. Flores, V. Menon, G. H. Glover, H. B. Solvason, H. Kenna, A. L. Reiss, and A. F. Schatzberg, "Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus," *Biological psychiatry* **62**, 429 (2007).

[71] J. P. Ramirez-Mahaluf, J. Perramon, B. Otal, P. Villoslada, and A. Compte, "Subgenual anterior cingulate cortex controls sadness-induced modulations of cognitive and emotional network hubs," *Scientific reports* **8**, 8566 (2018).

[72] A. V. Utevsky, D. V. Smith, and S. A. Huettel, "Precuneus is a functional core of the default-mode network," *Journal of Neuroscience* **34**, 932 (2014).

[73] N. S. Werner, T. Meindl, J. Materne, R. R. Engel, D. Huber, M. Riedel, M. Reiser, and K. Hennig-Fast, "Functional mri study of memory-related brain regions in patients with depressive disorder," *Journal of affective disorders* **119**, 124 (2009).

[74] E. R. Kenny, J. T. O'brien, D. A. Cousins, J. Richardson, A. J. Thomas, M. J. Firbank, and A. M. Blamire, "Functional connectivity in late-life depression using resting-state functional magnetic resonance imaging," *The American Journal of Geriatric Psychiatry* **18**, 643 (2010).

[75] S. L. Naismith, J. Lagopoulos, P. B. Ward, C. G. Davey, C. Little, and I. B. Hickie, "Fronto-striatal correlates of impaired implicit sequence learning in major depression: an fmri study," *Journal of affective disorders* **125**, 256 (2010).

[76] A. Dutta, S. McKie, and J. W. Deakin, "Resting state networks in major depressive disorder," *Psychiatry Research: Neuroimaging* **224**, 139 (2014).

[77] B. Fischl, "Freesurfer," *Neuroimage* **62**, 774 (2012).

[78] W. Truong, L. Minuzzi, C. N. Soares, B. N. Frey, A. C. Evans, G. M. MacQueen, and G. B. Hall, "Changes in cortical thickness across the lifespan in major depressive disorder," *Psychiatry Research: Neuroimaging* **214**, 204 (2013).

[79] R. S. Mackin, D. Tosun, S. G. Mueller, J.-Y. Lee, P. Insel, N. Schuff, D. Truran-Sacrey, P. Arean, J. C. Nelson, and M. W. Weiner, "Patterns of reduced cortical thickness in late-life depression and relationship to psychotherapeutic response," *The American Journal of Geriatric Psychiatry* **21**, 794 (2013).

[80] M. Niu, Y. Wang, Y. Jia, J. Wang, S. Zhong, J. Lin, Y. Sun, L. Zhao, X. Liu, L. Huang, *et al.*, "Common and specific abnormalities in cortical thickness in patients with major depressive and bipolar disorders," *EBioMedicine* **16**, 162 (2017).

[81] P. R. Escalona, B. Early, W. M. McDonald, P. M. Doraiswamy, S. A. Shah, M. M. Husain, O. B. Boyko, G. S. Figiel, E. H. Ellinwood Jr, C. B. Nemeroff, *et al.*, "Reduction of cerebellar volume in major depression: a controlled mri study," *Depression* **1**, 156 (1993).

[82] T. Moberget, D. Alnaes, T. Kaufmann, N. T. Doan, A. Cordova-Palomera, L. B. Norbom, J. Rokicki, D. van der Meer, O. A. Andreassen, and L. T. Westlye, "Cerebellar grey matter volume is associated with cognitive function and psychopathology in adolescence," *Biological psychiatry* (2019).

[83] L. Baldaçara, J. G. F. Borgio, A. L. T. d. Lacerda, and A. P. Jackowski, "Cerebellum and psychiatric disorders," *Brazilian Journal of Psychiatry* **30**, 281 (2008).

[84] K. R. R. Krishnan, W. M. McDonald, P. R. Escalona, P. M. Doraiswamy, C. Na, M. M. Husain, G. S. Figiel, O. B. Boyko, E. H. Ellinwood, and C. B. Nemeroff, "Magnetic resonance imaging of the caudate nuclei in depression: preliminary observations," *Archives of general psychiatry* **49**, 553 (1992).

[85] M. J. Kim, J. P. Hamilton, and I. H. Gotlib, "Reduced caudate gray matter volume in women with major depressive disorder," *Psychiatry Research: Neuroimaging* **164**, 114 (2008).

[86] D. Robinson, H. Wu, R. A. Munne, M. Ashtari, J. M. J. Alvir, G. Lerner, A. Koreen, K. Cole, and B. Bogerts, "Reduced caudate nucleus volume in obsessive-compulsive disorder," *Archives of General Psychiatry* **52**, 393 (1995).

[87] E. Bora, B. Harrison, C. Davey, M. Yücel, and C. Pantelis, "Meta-analysis of volumetric abnormalities in cortico-striatal-pallidal-thalamic circuits in major depressive disorder," *Psychological medicine* **42**, 671 (2012).

[88] R. Batuwita and V. Palade, "Class imbalance learning methods for support vector machines," (2013).

[89] N. Japkowicz, in *Proc. of the Int'l Conf. on Artificial Intelligence* (2000).

Appendices

Appendix A

MRI parameters of the fMRI data

\\USER\Hersenen Research 3T\Psychiatrie\rTMS\RestingState

TA: 10:12 PAT: 3 Voxel size: 3.0×3.0×3.0 mm Rel. SNR: 1.00 SIEMENS: ep2d_pace

Properties

Prio Recon	Off
Before measurement	
After measurement	
Load to viewer	On
Inline movie	Off
Auto store images	On
Load to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Start measurement without further preparation	On
Wait for user to start	On
Start measurements	single

Routine

Slice group 1	
Slices	40
Dist. factor	10 %
Position	R10.6 A3.1 H21.9
Orientation	T > C-13.0 > S4.9
Phase enc. dir.	A >> P
Rotation	0.00 deg
Phase oversampling	0 %
FoV read	192 mm
FoV phase	100.0 %
Slice thickness	3.0 mm
TR	2000 ms
TE	29.0 ms
Averages	1
Concatenations	1
Filter	None
Coil elements	HEA;HEP

Contrast

MTC	Off
Flip angle	90 deg
Fat suppr.	Fat sat.
Averaging mode	Long term
Reconstruction	Magnitude
Measurements	300
Delay in TR	0 ms
Multiple series	Off

Resolution

Base resolution	64
Phase resolution	100 %
Phase partial Fourier	Off
Interpolation	Off
PAT mode	GRAPPA
Accel. factor PE	3
Ref. lines PE	36
Matrix Coil Mode	Auto (Triple)
Reference scan mode	Separate
Distortion Corr.	Off
Prescan Normalize	Off
Raw filter	On
Elliptical filter	Off
Hamming	Off

Geometry

Multi-slice mode	Interleaved
Series	Interleaved
Special sat.	None

System

Body	Off
HEP	On
HEA	On
Positioning mode	REF
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Sum of Squares
AutoAlign	Head > Brain Atlas
Auto Coil Select	Default
Shim mode	Standard
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Silicone	Off
? Ref. amplitude 1H	0.000 V
Adjustment Tolerance	Auto
Adjust volume	
Position	R10.6 A3.1 H21.9
Orientation	T > C-13.0 > S4.9
Rotation	0.00 deg
R >> L	192 mm
A >> P	192 mm
F >> H	132 mm

Physio

1st Signal/Mode	None

BOLD

GLM Statistics	Off
Dynamic t-maps	Off
Starting ignore meas	0
Ignore after transition	0
Model transition states	On
Temp. highpass filter	On
Threshold	4.00
Paradigm size	20
Meas[1]	Baseline
Meas[2]	Baseline
Meas[3]	Baseline
Meas[4]	Baseline
Meas[5]	Baseline
Meas[6]	Baseline
Meas[7]	Baseline
Meas[8]	Baseline
Meas[9]	Baseline
Meas[10]	Baseline
Meas[11]	Active
Meas[12]	Active
Meas[13]	Active
Meas[14]	Active
Meas[15]	Active
Meas[16]	Active
Meas[17]	Active
Meas[18]	Active
Meas[19]	Active
Meas[20]	Active
Motion correction	Off
Spatial filter	Off

Sequence

Introduction	On
Bandwidth	2694 Hz/Px
Free echo spacing	Off
Echo spacing	0.48 ms
EPI factor	64
RF pulse type	Normal
Gradient mode	Fast*